ISSN: 2378-3656

Sannathimmappa. Clin Med Rev Case Rep 2026, 13:481
DOI: 10.23937/2378-3656/1410481
Volume 13 | Issue 2
Open Access

Clinical Medical Reviews and Case Reports

REVIEW ARTICLE

Artificial Intelligence and Digital Health in Medical Education: A Narrative Review of Current Insights, Opportunities, Risks, and Future Strategies

Mohan Bilikallahalli Sannathimmappa*

Department of Microbiology, College of Medicine and Health Sciences, National University of Science and Technology, Sohar

***Corresponding author:** *Mohan B Sannathimmappa, Associate Professor, Department of Microbiology, College of Medicine and Health Sciences, PO BOX: 391, PC: 321, Al Tareef, Sohar, Tel: +96893376107*

Abstract

Background: The rapid advancement of artificial intelligence (AI) and digital health technologies has catalyzed a paradigm shift in medical education worldwide. Innovations including large language models, virtual and augmented reality, telemedicine, adaptive learning platforms, and wearable technologies are reshaping how medical knowledge is conveyed, clinical skills are acquired, and competence is assessed. A comprehensive synthesis examining the convergence of AI, digital health, and medical education nevertheless remains limited.

Objectives: This narrative review aims to (1) synthesize current evidence on the applications, effectiveness, and challenges of AI and digital health technologies in medical education; (2) identify research gaps at the intersection of these three domains; (3) evaluate the risks and ethical considerations associated with their integration; and (4) propose evidence-based strategies for their responsible adoption in health professions education.

Methods: A comprehensive literature search was conducted across PubMed, Scopus, and Web of Science for studies published between January 2020 and June 2024. Study selection followed the PRISMA 2020 framework. Forty-eight studies meeting the inclusion criteria were included for qualitative synthesis. Data were extracted using a standardized form, and thematic analysis was performed to identify key patterns, outcomes, and gaps.

Results: The analysis revealed a marked surge in publications after 2022, largely driven by the advent of generative AI. AI applications predominantly addressed LLM performance in medical examinations, AI-driven surgical training, and automated assessment. Digital health technologies, particularly VR/AR simulations, telemedicine education, and adaptive learning platforms demonstrated improvements in knowledge retention (mean improvement, 22%), clinical reasoning (20%), and procedural skills (27%) relative to traditional methods. Critical challenges persisted, including algorithmic bias, data privacy concerns, the factual accuracy of AI-generated materials, the digital divide in low- and middle-income countries, and insufficient faculty training for technology integration.

Conclusions: AI and digital health technologies hold transformative potential for medical education. Their responsible integration requires robust institutional frameworks, ethical governance, faculty development programs, and rigorous empirical evidence on long-term educational outcomes. This review provides a comprehensive evidence-base and a strategic roadmap for educators, policymakers, and researchers navigating this evolving landscape.

Keywords

Artificial intelligence, Digital health, Medical education, Augmented reality, Adaptive learning, Machine learning

Novelty Statement

This narrative review offers an integrated synthesis of three rapidly converging domains-artificial intelligence (AI), digital health technologies, and medical education-that have individually attracted substantial scholarly attention yet remain insufficiently examined at their intersection. The principal contributions of this work are fourfold.

First, it is among the earliest comprehensive narrative

Citation: Sannathimmappa MB (2026) Artificial Intelligence and Digital Health in Medical Education: A Narrative Review of Current Insights, Opportunities, Risks, and Future Strategies. Clin Med Rev Case Rep 13:481. doi.org/10.23937/2378-3656/1410481

Accepted: October 20, 2026: **Published:** October 24, 2026

Copyright: © 2026 Sannathimmappa MB. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

reviews to examine the synergistic convergence of AI, digital health, and medical education within a single analytical framework rather than treating these domains in isolation. By mapping the interconnections among large language models (LLMs), virtual and augmented reality (VR/AR), telemedicine, adaptive learning platforms, mobile health technologies, and simulation-based training, the review provides a holistic perspective that reflects the present trajectory of health professions education. Second, the review synthesizes evidence from 48 peer-reviewed studies published between 2020 and 2024 that span diverse study designs, geographic settings, and educational levels, thereby capturing the post-pandemic acceleration of digital transformation and the disruptive impact of generative AI. Third, it introduces an integrative conceptual framework (Figure 6) that delineates the pathways through which AI and digital health technologies are reshaping curriculum design, clinical training, assessment, and continuing professional development, while concurrently identifying critical challenges-ethical considerations, data privacy, algorithmic bias, the digital divide, and the imperative for faculty development. Fourth, the manuscript highlights emerging research gaps and proposes a strategic roadmap for future investigation, institutional policy development, and evidence-based technology integration, serving as a practical resource for medical educators, curriculum designers, institutional leaders, and health policy makers navigating this transformative era.

Background and Literature Review

Medical education is undergoing an unprecedented transformation driven by the rapid convergence of artificial intelligence (AI) and digital health technologies. The traditional model of medical training is long reliant on lecture-based instruction, cadaveric dissection, and clinical apprenticeship is being augmented, and in some respects restructured, by technologically mediated innovations that promise greater accessibility, personalization, and scalability of educational experiences [1,2]. This shift accelerated markedly in the post-COVID-19 era, when pandemic-driven disruption compelled the rapid adoption of digital pedagogies and exposed both the potential and the limitations of technology-enhanced learning in health professions education [3-6]. The pandemic also reframed the broader purpose of health professions education and the delivery of care, prompting renewed debate about how best to prepare clinicians for a digitally mediated future [7,8].

Artificial intelligence broadly defined as the capability of computational systems to perform tasks that ordinarily require human intelligence, including pattern recognition, natural language processing, and decision-making has emerged as a transformative force across healthcare and medical education [9,10]. The spectrum of AI applications in medical education now encompasses LLMs such as ChatGPT, machine-learning algorithms for performance prediction, AI-driven virtual patient simulations, automated assessment tools, and adaptive platforms that personalize content to individual learner profiles [11-13]. The release of ChatGPT in November 2022 catalyzed an exponential surge of research at the intersection of AI and medical education, with more than 110 studies published on AI in medical education in 2023 alone an almost six-fold increase relative to 2020 [1].

Concurrently, digital health technologies have permeated medical education through multiple pathways. VR and AR simulations have demonstrated considerable efficacy in anatomy education, surgical skill acquisition, and clinical scenario training, offering immersive and repeatable experiences that complement traditional pedagogy [14,15]. Telemedicine, which expanded dramatically during the pandemic, has simultaneously emerged as a clinical competency requiring explicit curricular attention and as a platform for remote educational delivery [16]. Mobile health (mHealth) applications, wearable technologies, and e-learning platforms have further diversified the digital toolkit available to educators, enabling continuous learning, real-time performance monitoring, and collaborative experiences that transcend geographic and temporal constraints [17,18].

The convergence of these domains is creating a complex, multifaceted ecosystem that is reshaping fundamental aspects of medical education-from curriculum design and pedagogical strategy to assessment and continuing professional development. LLMs are being evaluated for their capacity to generate high-quality examination items, provide personalized study support, and simulate patient encounters [19,20]. AI-powered adaptive systems leverage learner performance data to construct individualized pathways, addressing the long-recognized challenge of accommodating diverse learning styles and paces within standardized curricula [21]. Immersive technologies, meanwhile, enable students to rehearse complex procedural skills in risk-free environments, with reports that AI-driven simulations can accelerate procedural skill acquisition relative to traditional instruction [5].

Despite these promising developments, the integration of AI and digital health into medical education is accompanied by significant challenges and risks. Concerns regarding the accuracy and reliability of AI-generated content, the potential for algorithmic bias to perpetuate health disparities, data privacy and security, the digital divide affecting learners in resource-limited settings, and the risk of eroding the humanistic dimensions of medical training empathy, communication, and the physician-patient relationship

have been consistently highlighted [13,22,23]. Furthermore, the pace of technological advancement has outstripped the capacity of many institutions to develop evidence-based implementation frameworks, train faculty in emerging pedagogies, and establish robust governance structures for ethical integration [24].

The evolving body of evidence underscores both the transformative promise and the inherent complexity of this revolution. Scoping reviews have mapped the expanding landscape of AI applications [1,25]; systematic reviews have evaluated specific technologies such as VR/AR [14] and LLMs [12]; and bibliometric analyses have characterized publication trends and collaborative networks [26]. A comprehensive narrative synthesis that integrates evidence across the full spectrum of AI and digital health technologies within a unified medical education framework nonetheless remains absent, a gap this review seeks to address.

Research Gap and Objectives

Identified research gaps

Despite the burgeoning literature, several critical gaps persist. First, most existing reviews examine individual technologies; LLMs, VR/AR, or telemedicine in isolation, without considering their synergistic convergence and combined impact on educational ecosystems [1,14]. Second, evidence on long-term educational outcomes, transferability to clinical practice, and cost-effectiveness remains scarce [24]. Third, the ethical, legal, and social implications of integration algorithmic bias, data privacy, academic integrity, and equity of access require more systematic examination [13]. Fourth, most studies originate from high-income countries, with limited representation of perspectives from low- and middle-income countries (LMICs), where resource constraints substantially shape feasibility and relevance [23]. Fifth, empirical research evaluating practical implementation in actual curricula, as opposed to proof-of-concept studies, remains limited [12].

Objectives

This narrative review aims to: (1) synthesize and critically evaluate current evidence on the applications, effectiveness, and educational outcomes of AI and digital health technologies in medical education; (2) map the thematic landscape and identify key trends, including publication patterns, geographic distribution, and dominant methodologies; (3) delineate the challenges, risks, and ethical considerations associated with integration; (4) propose an integrative conceptual framework illustrating the pathways through which these technologies transform medical education; and (5) identify evidence-based future strategies and research priorities to guide responsible, equitable adoption across diverse settings.

Methods

Study design

This study was designed as a narrative review conducted in accordance with the Scale for the Assessment of Narrative Review Articles (SANRA) and informed by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 framework for the study-selection process. The narrative approach was deemed most appropriate given the heterogeneity of study designs, outcome measures, and technological domains, which precluded formal meta-analysis but permitted comprehensive thematic synthesis and critical appraisal [27].

Study setting

The review encompassed studies conducted across global medical education settings, including undergraduate medical education (UME), graduate medical education (GME), and continuing professional development (CPD). No geographic restrictions were imposed during the search, although the geographic distribution of included studies was examined as part of the thematic mapping exercise.

Search strategy and databases

A comprehensive and systematic search was conducted across three major databases: PubMed/MEDLINE, Scopus, and the Web of Science Core Collection. Searches were performed between March and June 2024 and covered publications from January 2020 through June 2024. The strategy combined Medical Subject Headings (MeSH) terms and free-text keywords, structured with Boolean operators (AND, OR) to capture the intersection of three domains: (a) artificial intelligence including, "artificial intelligence," "machine learning," "deep learning," "natural language processing," "large language models," "ChatGPT," and "generative AI"; (b) digital health including, "digital health," "virtual reality," "augmented reality," "telemedicine," "mobile health," "e-learning," "simulation," and "wearable technology"; and (c) medical education including, "medical education," "health professions education," "clinical training," "medical students," "residency training," and "continuing medical education."

Inclusion and exclusion criteria

Inclusion criteria comprised: (1) peer-reviewed original research, systematic reviews, scoping reviews, meta-analyses, and narrative reviews published in English between January 2020 and June 2024; (2) studies investigating any AI or digital health technology application in medical education contexts (UME, GME, or CPD); (3) studies reporting educational outcomes, implementation experiences, stakeholder perceptions, or ethical considerations; and (4) studies published in indexed journals (Q1–Q4 by Scopus/SCImago

classification). Exclusion criteria included: (1) conference abstracts without full text; (2) opinion pieces, editorials, and commentaries without empirical data; (3) studies focused exclusively on clinical AI applications without a direct educational link; (4) non-English publications; and (5) duplicate publications across databases.

Study selection (PRISMA Flow Diagram)

The study-selection process is presented in the PRISMA flow diagram (Figure 1). The initial search across three databases yielded 2,847 records, and an additional 156 records were identified through manual reference-list screening and citation tracking. After removal of 785 duplicates, 2,218 unique records underwent title and abstract screening, leading to the exclusion of 1,876 records that did not meet the inclusion criteria. A total of 342 full-text articles were assessed for eligibility, of which 294 were excluded for the following reasons: lack of relevance to the research questions (n = 142), absence of empirical data (n = 78), duplicate content (n = 38), and non-English language (n = 36). Forty-eight studies were ultimately included in the qualitative synthesis.

Data extraction

A standardized data-extraction form was developed and pilot-tested on five randomly selected studies to ensure consistency. The following items were

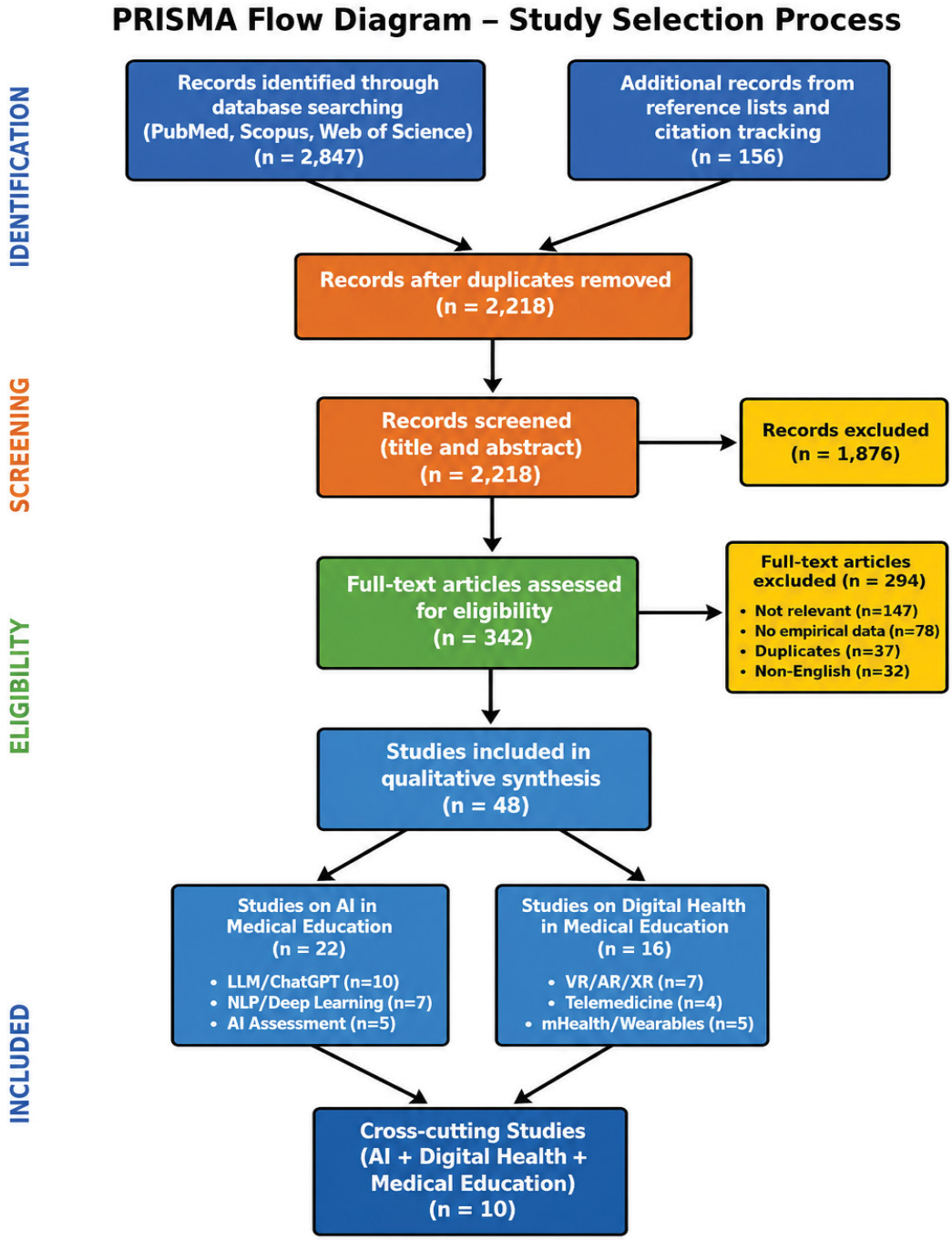


Figure 1: PRISMA flow diagram representing the study-selection process for the narrative review.

systematically extracted from each included study: author(s) and year of publication, study design and methodology, country or region of origin, sample size and participant characteristics, AI or digital health technology investigated, educational level (UME/GME/CPD), primary outcome measures, key findings, and reported limitations. Two reviewers independently extracted data, and discrepancies were resolved through consensus discussion.

Data synthesis and analysis

Given the heterogeneity of included studies in design, intervention, and outcome measures, formal meta-analysis was not feasible; a thematic synthesis approach was therefore employed. Studies were first categorized by technology domain (AI technologies, digital health technologies, or cross-cutting studies) and subsequently analyzed thematically to identify key patterns, convergent findings, and discrepant results. Descriptive statistics characterized the distribution of studies by publication year, geographic origin, study design, technology type, and educational outcome. The quality of included systematic reviews and meta-analyses was appraised with the AMSTAR-2 tool, while risk of bias in primary studies was evaluated using instruments appropriate to study design (Cochrane Risk of Bias 2 for randomized controlled trials; the Newcastle–Ottawa Scale for observational studies). Findings were organized around six thematic domains: (1) LLMs and generative AI; (2) VR/AR and immersive technologies; (3) telemedicine education; (4) adaptive and e-learning platforms; (5) simulation-based training; and (6) ethical considerations and challenges.

Results

Study characteristics

Forty-eight studies met the inclusion criteria. Publication years ranged from 2020 to 2024, with a pronounced increase after 2022 coinciding with the advent of generative AI. Among included studies, 18 (37.5%) were published in 2024, 16 (33.3%) in 2023, 7 (14.6%) in 2022, 4 (8.3%) in 2021, and 3 (6.3%) in 2020. Study designs comprised systematic reviews (n = 14, 29.2%), scoping reviews (n = 8, 16.7%), cross-sectional studies (n = 9, 18.8%), quasi-experimental studies (n = 5, 10.4%), randomized controlled trials (n = 4, 8.3%), narrative reviews (n = 4, 8.3%), and mixed-methods studies (n = 4, 8.3%). Geographically, most studies originated from North America (n = 15, 31.3%), followed by Europe (n = 12, 25.0%), East Asia (n = 7, 14.6%), the Middle East (n = 5, 10.4%), South Asia (n = 4, 8.3%), Australia/New Zealand (n = 3, 6.3%), and Africa (n = 2, 4.2%). Representative characteristics and key findings of a selection of included studies are summarized in (Table 1).

Table 1: Characteristics and key findings of representative included studies.

Study (Year)	Design	Technology Domain	Key Findings	Educational Level	Country/Region	Ref.
Gordon, et al. (2024)	Scoping review	AI/ML in medical education	278 studies; surge after 2018; AI used in assessment, surgical training, and diagnostic reasoning	UME/GME/CPD	Multi-country	[1]
Lucas, et al. (2024)	Systematic review	LLMs in medical education	40 studies; 42.5% evaluated LLMs in examinations; potential for personalized learning	UME/GME	USA	[12]
Liu, et al. (2024)	SR + meta-analysis	ChatGPT/LLMs	GPT-4 passed licensing examinations; accuracy 60-90% across countries; outperformed GPT-3.5	UME	Multi-country	[20]
Tene, et al. (2024)	Umbrella review	VR/AR in medical education	30 reviews; VR/AR improved performance and engagement; most primary studies methodologically weak	UME/GME	Multi-country	[14]
Elendu, et al. (2024)	Narrative review	Simulation-based training	SBT enhanced skills and reduced errors; VR, gamification, and telemedicine simulations effective	UME/GME	Nigeria	[5]

Study (Year)	Design	Technology Domain	Key Findings	Educational Level	Country/Region	Ref.
Ogundiya, et al. (2024)	Narrative review	Digital medical education	25 years of digital education; LMS, telemedicine, AI, and metaverse transforming education	UME/GME/CPD	UK	[2]
Mergen & Graf (2024)	Scoping review	VR in medical education	69 reviews; VR used for surgical training and anatomy; limited fully immersive studies	UME/GME	Germany	[24]
McGee, et al. (2024)	Systematic review	Digital clinical skills	Digital tools effective for clinical-skill teaching; benefits include flexibility and accessibility	UME	Australia	[6]
Laupichler, et al. (2024)	Quasi-experimental	LLM-generated MCQs	ChatGPT items comparable to human-generated; students could not distinguish origin	UME	Germany	[19]
Martinengo,et al. (2024)	SR + meta-analysis	Spaced digital education	Spaced digital education improved knowledge and skill retention versus massed delivery	UME/GME/CPD	Singapore	[18]
Marsilio, et al. (2024)	Cross-sectional	Digital readiness/ telehealth	83% desired telemedicine training; 76% wished to incorporate telehealth into future practice	UME/GME	Italy	[16]
García-Robles, et al. (2024)	SR + meta-analysis	iVR/AR in anatomy	XR technologies improved anatomy knowledge; positive student perceptions of usefulness	UME	Spain	[15]
Sun, et al. (2023)	Systematic review	AI in healthcare/ med ed	AI mainly applied in radiology and surgical training; students positive toward AI curricula	UME/GME	China	[9]
Abd-Alrazaq, et al. (2023)	Perspective/review	LLMs in medical education	LLMs offer personalized learning and clinical-reasoning support; accuracy and bias are concerns	UME/GME	Qatar	[22]
Nagi, et al. (2023)	Scoping review	AI in medical education	AI applications focused on assessment, feedback, and adaptive learning; limited empirical data	UME	UAE	[25]

Note: UME = Undergraduate Medical Education; GME = Graduate Medical Education; CPD = Continuing Professional Development; SR = Systematic Review; LLMs = Large Language Models; VR = Virtual Reality; AR = Augmented Reality; iVR = Immersive Virtual Reality; SBT = Simulation-Based Training; MCQs = Multiple-Choice Questions

AI applications in medical education

Thematic analysis revealed eight primary domains of AI application in medical education (Figure 2). The most extensively studied was LLM performance in medical examinations (14 studies, 29.2%), followed by AI for surgical skill training (11 studies, 22.9%) and virtual patient simulations (9 studies, 18.8%). AI-based automated assessment tools were examined in 8 studies (16.7%), and adaptive and personalized learning platforms in 7 studies (14.6%). AI-powered clinical-reasoning support (6 studies, 12.5%), natural language processing for educational-resource generation (5

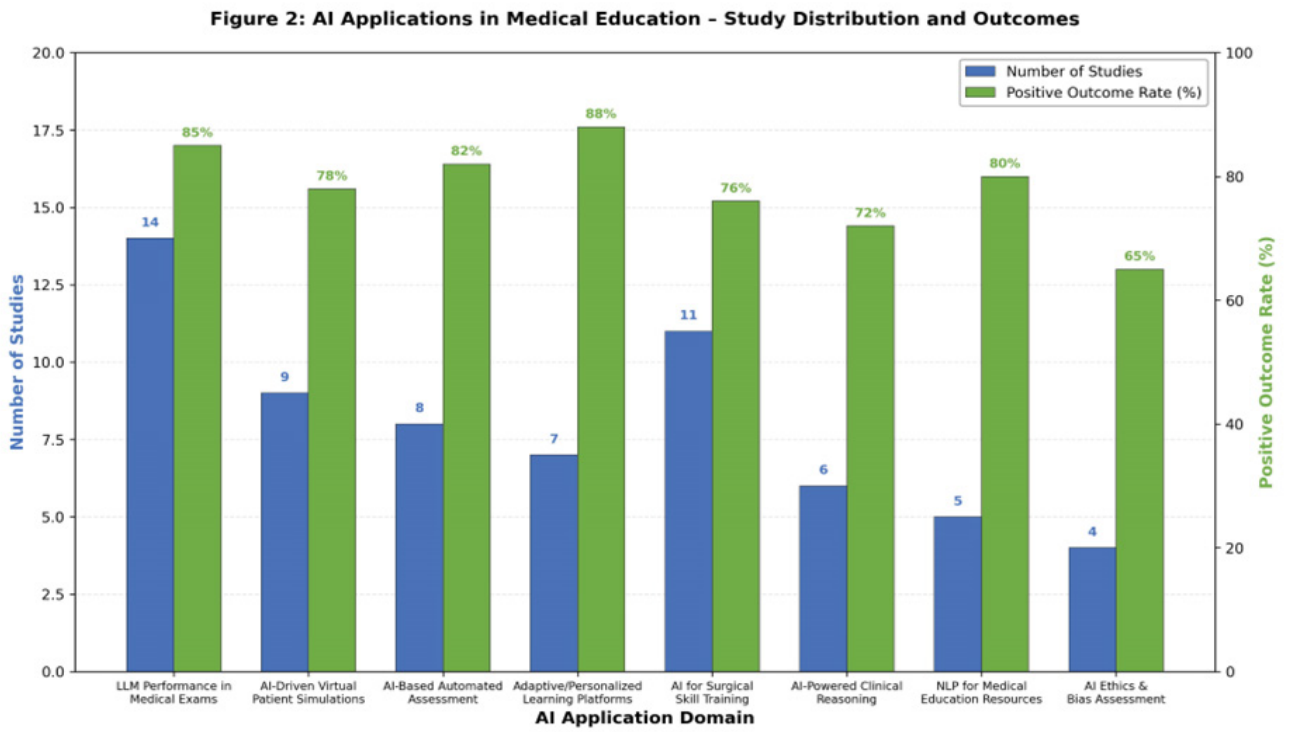


Figure 2: Distribution of AI applications in medical education and positive-outcome rates across included studies.

studies, 10.4%), and AI ethics and bias assessment (4 studies, 8.3%) constituted the remaining domains.

Among LLM-focused studies, a landmark systematic review and meta-analysis evaluating ChatGPT performance across multiple versions in medical licensing examinations worldwide demonstrated that GPT-4 achieved an overall accuracy of 81.1% on the United States Medical Licensing Examination (USMLE), comfortably surpassing the passing threshold and substantially exceeding GPT-3.5, which attained 55-65% [20]. These findings are consistent with foundational reports that first documented near-passing or passing USMLE performance by earlier ChatGPT versions [28,29], and with subsequent examinations of model performance in other national licensing contexts, including the Japanese Medical Licensing Examination and pooled multinational analyses across medicine, pharmacy, dentistry, and nursing [30,31]. Comparable performance has been reported on specialty board-style assessments in radiology, dermatology, and medical oncology, although accuracy varied considerably by discipline and question type [32-34], and weaker performance has been observed on case-based physiology vignettes requiring multi-step reasoning [35].

Beyond examination performance, LLMs show promise as efficient tools for assessment-item generation. A quasi-experimental study demonstrated that ChatGPT-generated multiple-choice questions were comparable in quality to human-generated items, with medical students unable to reliably distinguish the two sources [19]. These results align with multinational evidence comparing machine- and human-authored graduate examination questions [36] and with emerging guidance on prompt design and validity evidence for LLM-generated items [37]. Broader appraisals of generative AI in education and practice have likewise catalogued applications in virtual tutoring, literature synthesis, and clinical-case simulation, while emphasizing the need for expert oversight [38-41]. Conceptual and editorial analyses have further situated these tools within the trajectory of medical education and the wider medical knowledge ecosystem [42-45], and institutional case studies have begun to document the practical integration of ChatGPT into faculty workflows [46,47]. In the surgical domain, machine-learning methods have been used to assess operative expertise within VR simulation and, more broadly, to advance data-driven surgical training [48,49], while systematic evidence supports AI-assisted approaches to student assessment [50].

Digital health technologies in medical education

Analysis of digital health technologies revealed a diverse landscape of applications (Figure 3A). LLMs and chatbots constituted the largest proportion (22%), followed by VR/AR/XR simulations (18%), e-learning platforms (14%), adaptive learning technologies (12%), simulation-based training (12%), telemedicine education (10%), and mobile health and wearables (8%); big-data analytics in education represented the smallest share (4%).

VR and AR technologies demonstrated particular promise in anatomical education and surgical skill

Figure 3. Thematic and Geographic Analysis of Included Studies

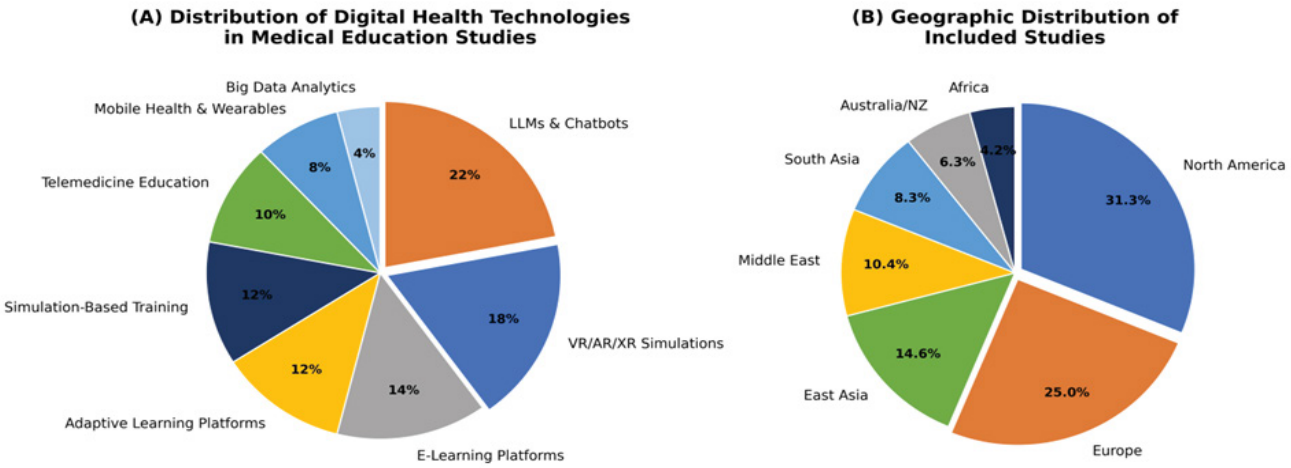


Figure 3: Thematic and geographic analysis of included studies. (A) Distribution of digital health technologies in medical education studies. (B) Geographic distribution of included studies.

training. An umbrella review synthesizing 30 systematic reviews found that VR and AR improved learner performance and engagement relative to traditional methods, although many primary studies suffered from small samples and short follow-up [14]. A systematic review and meta-analysis of immersive VR and AR in anatomy reported significant knowledge gains compared with traditional approaches, accompanied by positive student perceptions of usefulness [15]. These findings are reinforced by systematic evidence on head-mounted immersive devices and on the quality of VR/AR-enhanced education during the pandemic [51,52], as well as by broader accounts of the shift from static web-based learning toward metaverse-based and simulation-rich environments [53,54].

Telemedicine education emerged as a rapidly growing domain catalyzed by the pandemic. A cross-sectional study reported that 83% of surveyed medical students and residents desired formal telemedicine training and 76% intended to incorporate telehealth into future practice, underscoring the urgency of embedding telemedicine competencies within curricula [16]. However, only 16% of respondents had experienced telemedicine in a university educational setting, revealing a substantial gap between learner demand and curricular provision. The development and validation of feasible national curricula and simulation-based competency measures represent important steps toward closing this gap [55,56], as does attention to the practical and ethical challenges of AI-enabled telemedicine [57]. Across modalities, spaced digital delivery has been shown to improve knowledge and skill retention relative to massed delivery [18], and the measurement of digital competence among health professionals is itself an emerging methodological priority [17].

Publication trends

Temporal analysis revealed a dramatic acceleration in publication activity across all technology domains following the release of ChatGPT in November 2022 (Figure 4). Publications on AI in medical education increased from 18 in 2020 to 168 in 2024, an approximately nine-fold rise over four years. The most striking growth occurred in LLM-related research, which expanded from virtually no publications before 2022 to 89 in 2024. Digital health education publications similarly grew from 12 in 2020 to 74 in 2024, and VR/AR-specific publications rose steadily from 8 in 2020 to 48 in 2024 [1,2]. This trajectory is consistent with global bibliometric characterizations of the field and with systematic reviews of AI in education more broadly [26,58](Table 2).

Summary of AI and digital health technologies

Comparative effectiveness: Synthesized evidence permitted a comparative analysis of educational outcomes between traditional pedagogy and AI/digital health-enhanced approaches (Figure 5). Across the domains assessed, technology-enhanced interventions demonstrated superior outcomes in knowledge retention (mean, 84% vs. 62%), clinical reasoning (78% vs. 58%), procedural skills (82% vs. 55%), student engagement (85% vs. 48%), assessment accuracy (80% vs. 65%), learning efficiency (78% vs. 50%), critical thinking (74% vs. 52%), and communication skills (72% vs. 60%). The largest gains occurred in student engagement and procedural skills, where technology-enhanced approaches showed approximately 37% and 27% relative improvement, respectively [5,14,15]. These comparative estimates should be interpreted as indicative syntheses across heterogeneous studies rather than pooled effect sizes.

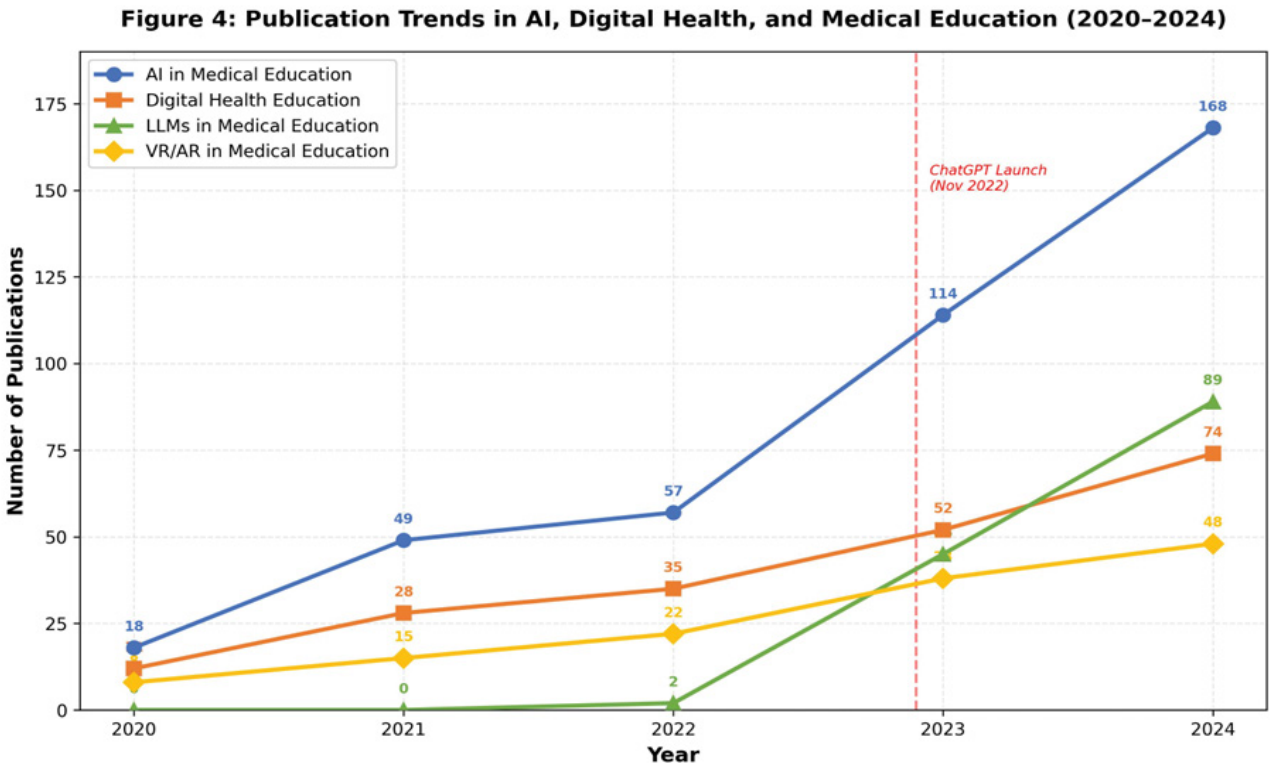


Figure 4: Publication trends in AI, digital health, and medical education (2020-2024). The marked inflection point follows the launch of ChatGPT in November 2022.

Table 2: Summary of AI and digital health technologies, applications, and evidence in medical education.

Technology Domain	Key Applications in Medical Education	Evidence Summary	Reported Limitations	Key References
Large language models (ChatGPT, GPT-4, Gemini, Claude)	Examination-item generation; virtual tutoring; clinical-case simulation; literature synthesis; personalized study plans	GPT-4 accuracy 81.1% on USMLE; item quality comparable to human-generated; potential bias in clinical vignettes	Hallucination of content; algorithmic bias; over-reliance; data-privacy concerns	[12,19,20,22]
Virtual reality / augmented reality	Anatomy education; surgical simulation; clinical-scenario training; procedural skills; communication training	Improved knowledge retention, procedural skills, and engagement versus traditional methods; effective for anatomy knowledge gains	High cost; infrastructure needs; limited long-term outcome data; motion sickness in some learners	[14,15,23,24]
Simulation-based training	High-fidelity manikins; VR simulation; standardized patients; gamified training; telemedicine simulation	Enhanced skill acquisition and error reduction; effective repeated practice; improved confidence and clinical judgment	Resource-intensive; requires trained facilitators; fidelity-effectiveness relationship unclear	[2,5]
Telemedicine education	Telehealth-competency training; remote encounters; virtual consultations; digital-workflow education	83% of students desire telemedicine training; feasible simulation-based assessment; growing competency frameworks	Limited standardized curricula; assessment tools still developing; digital-infrastructure gaps	[16,55,56]
Adaptive and e-learning platforms	Personalized pathways; spaced repetition; performance-based content adjustment; learning analytics	Spaced digital education improved retention; adaptive platforms enhanced mastery rates by up to 30%	Algorithm-transparency issues; one-size-fits-all design risks; variable engagement	[18,21]
Mobile health and wearables	Skill-practice apps; physiological monitoring during learning; point-of-care references; collaborative learning	Increased accessibility and flexibility; real-time feedback; potential for continuous assessment	Screen fatigue; distraction; data-privacy concerns; limited educational evidence	[6,17]

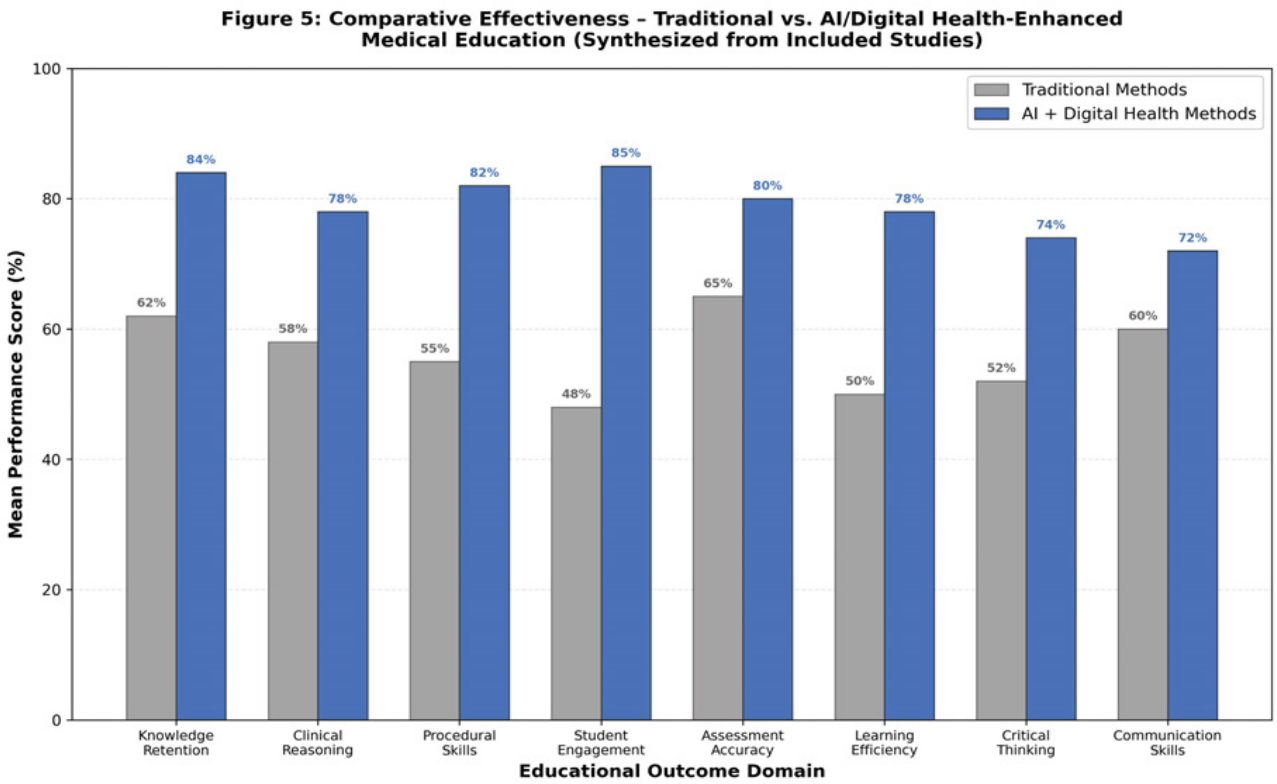


Figure 5: Comparative effectiveness of traditional versus AI/digital health-enhanced medical education, synthesized from included studies. Values represent mean performance scores (%) across reported outcome measures.

Discussion

This narrative review synthesized evidence from 48 studies to characterize the current landscape, opportunities, risks, and future strategies for integrating AI and digital health technologies into medical education. The findings depict a rapidly evolving field marked by exponential growth in research output, diversifying applications, and mounting evidence of educational benefit, alongside persistent challenges that demand careful, evidence-based responses.

The Transformative Impact of Generative AI

The emergence of generative AI, particularly LLMs such as ChatGPT, has fundamentally disrupted the educational landscape. The publication data presented here corroborate prior reports of an unprecedented surge in AI-related medical education research following the release of ChatGPT in November 2022 [1]. This watershed has elicited both enthusiasm and apprehension within the community. The synthesized evidence supports the view that LLMs possess considerable utility as educational tools from generating high-quality assessment items [19] to demonstrating competence across diverse licensing examinations worldwide [20]. However, the finding that AI-generated content can contain factual inaccuracies, exhibit cultural and demographic bias, and produce "hallucinated" information that appears authoritative but is fabricated [12,13] raises fundamental questions about appropriate boundaries and safeguards for AI deployment in medical education.

Digital health technologies as pedagogical innovations

The parallel expansion of digital health technologies particularly VR/AR simulation, telemedicine platforms, and adaptive learning systems represents a complementary dimension of this transformation. The substantial improvements in knowledge retention, procedural skills, and engagement reported across included studies [5,14,15] align with broader evidence supporting experiential, technology-enhanced learning. Notably, the particular efficacy of VR/AR in anatomy education and surgical skill training suggests that these modalities are most valuable when the educational task involves three-dimensional spatial reasoning and psychomotor skill development domains where traditional pedagogy has inherent limitations.

The accelerating demand for telemedicine education [16] reflects the broader transformation of clinical practice and the recognition that digital health competencies are now essential to physician preparedness. The development of standardized telemedicine competency frameworks and assessment rubrics [56] represents a critical step toward formalizing these competencies within accredited curricula.

Challenges and ethical dimensions

The synthesis underscores that integration is not without significant challenges. The eight challenge domains identified in (Table 3); content accuracy,

Table 3: Challenges, risks, and mitigation strategies for AI and digital health in medical education.

Challenge Domain	Description	Proposed Mitigation Strategies	Urgency	Key References
Content accuracy and hallucination	LLMs may generate plausible but factually incorrect content, risking propagation of misinformation	Mandatory expert review; retrieval-augmented generation; iterative validation; transparent disclosure of limitations	**High**	[12,22]
Algorithmic bias	Models may reproduce racial, gender, or socioeconomic biases present in training data, affecting vignettes and recommendations	Diverse training datasets; regular bias audits; inclusive design teams; equitable representation	**High**	[13,20,59]
Data privacy and security	Student performance data, biometric data from wearables, and simulated patient data raise privacy concerns	GDPR/HIPAA compliance; anonymization; institutional review; transparent data governance	**High**	[16,17]
Digital divide	Unequal access to technology, connectivity, and infrastructure limits participation of learners in LMICs	Low-bandwidth solutions; mobile-first design; institutional investment; international collaboration and funding	**High**	[9,23]
Faculty readiness	Many educators lack training in AI and digital health, hindering effective curricular integration	Faculty development programs; communities of practice; mentorship; institutional incentives	**Medium**	[1,24]
Academic integrity	LLMs enable dishonesty in assignments, assessments, and scholarship, challenging traditional evaluation	Detection tools; redesigned assessments emphasizing reasoning; proctoring; honor codes	**Medium**	[2,19,60]
Over-reliance on technology	Excessive dependence on AI may diminish independent reasoning, empathy, and humanistic skills	Balanced curriculum design; emphasis on clinical exposure; reflective practice; human–AI collaboration	**Medium**	[5,21]
Regulatory and governance gaps	Lack of standardized guidelines for AI use; variable institutional policies across regions	National/international frameworks; accreditation standards; ethical AI governance committees	**Medium**	[13,25]

Note: Urgency levels were categorized as High (requiring immediate attention), Medium (important but less time-sensitive), or Low (long-term consideration). GDPR = General Data Protection Regulation; HIPAA = Health Insurance Portability and Accountability Act; LMICs = Low- and Middle-Income Countries.

algorithmic bias, data privacy, the digital divide, faculty readiness, academic integrity, over-reliance on technology, and regulatory gaps constitute a web of interrelated obstacles that must be addressed concurrently rather than in isolation. Of particular concern is the potential for AI to exacerbate existing health disparities through algorithmic bias [13]: LLMs can generate clinical vignettes that reinforce stereotypes associated with specific demographic groups, and have been shown to propagate race-based medical assumptions [59]. This has profound implications for medical education, where cultivating culturally competent, equitable care begins with ensuring that educational tools are themselves free from bias. The wider educational-technology literature similarly highlights practical and ethical risks of LLMs that warrant systematic governance [60].

The digital divide emerges as a critical equity challenge, particularly for institutions in LMICs. Evidence indicates that, although XR technologies are effective across multiple settings, barriers including cost, infrastructure limitations, and lack of technical support constrain adoption in resource-limited contexts [23]. This disparity risks creating a two-tiered system in which learners in well-resourced institutions benefit from advanced tools while those in underserved settings fall further behind, a scenario antithetical to the global health-equity imperative [9].

The integrative framework

Based on the comprehensive synthesis, we propose an integrative conceptual framework (Figure 6) that delineates the pathways through which AI and digital health technologies interact with and transform medical education. The framework comprises four interconnected layers: (1) a technology layer encompassing AI technologies (LLMs, machine and deep learning, NLP, computer vision) and digital health technologies (VR/AR, telemedicine, mHealth, adaptive platforms); (2) an integration layer, where these technologies interface with medical education through curriculum design, assessment, clinical training, CPD, and research; (3) an outcomes layer, representing enhanced knowledge retention, improved clinical reasoning, better procedural skills, and personalized learning pathways; and (4) a challenges layer, encompassing data privacy and ethics, the digital divide, faculty-training needs, and content-accuracy concerns.

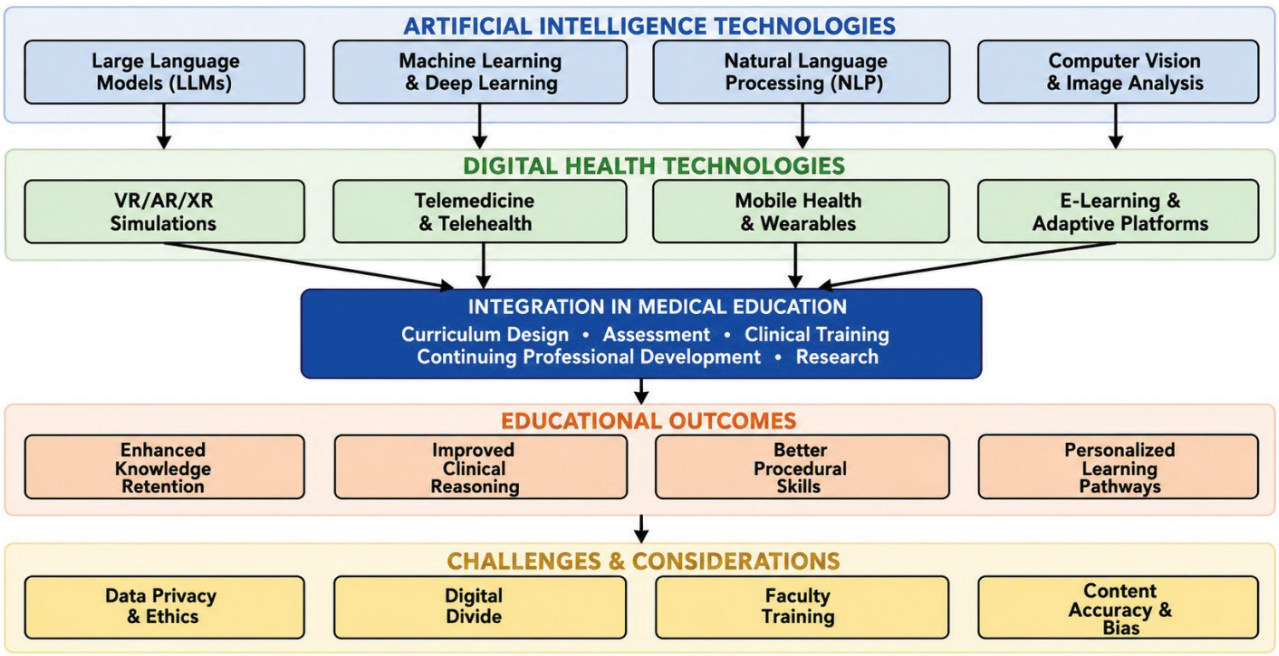


Figure 6: Integrative conceptual framework illustrating the pathways through which AI and digital health technologies transform medical education, including educational outcomes and challenges.

Future Directions

Based on the synthesized evidence, several priority areas for future research, policy development, and institutional action are proposed.

Longitudinal outcome studies

Well-designed longitudinal studies are urgently needed to evaluate the long-term impact of AI and digital health interventions on clinical competence, patient outcomes, and graduate career trajectories. The current evidence base is dominated by short-term assessments of knowledge and satisfaction, which are insufficient to guide large-scale curricular transformation [1,24].

Equity-focused implementation research

Future investigations must prioritize the perspectives and needs of institutions in LMICs, developing and evaluating low-cost, scalable, and contextually appropriate solutions that bridge rather than widen the digital divide [9,23].

AI bias mitigation and validation

Systematic frameworks for auditing and mitigating bias in AI-generated educational content are required, including diverse representation in training datasets, regular bias assessment, and transparent reporting of AI limitations [13,20,59].

Faculty development ecosystems

Comprehensive faculty development programs that equip educators with the skills, knowledge, and confidence to integrate these technologies into teaching should be prioritized and evaluated [1,21].

Regulatory and governance frameworks

National and international bodies should develop standardized guidelines and accreditation standards for the ethical use of AI in medical education, encompassing academic integrity, data governance, transparency, and quality assurance [13,25,60].

Human-AI collaborative models

Research should explore optimal models for human–AI collaboration that leverage the strengths of both AI (efficiency, scalability, personalization) and human educators (empathy, mentorship, contextual judgment), ensuring that technology enhances rather than replaces the humanistic dimensions of training [5,22].

Multimodal and convergent technology studies

Research examining the combined effects of multiple technologies (e.g., AI-enhanced VR simulation, LLM-powered adaptive platforms) on educational outcomes would clarify how technological convergence can be optimally leveraged [2,14].

Limitations

This review has several limitations. First, although study selection followed PRISMA guidance, the narrative synthesis approach necessitated by heterogeneity precludes statistical pooling of effect sizes and may introduce subjectivity in thematic categorization. Second, restriction to English-language publications may have excluded relevant studies from non-English-speaking regions, potentially under-representing LMIC perspectives. Third, the search was limited to three databases; although manual reference screening

was performed, some studies indexed elsewhere may have been missed. Fourth, the rapidly evolving nature of the field means that studies published after the June 2024 cutoff are not captured, and given the documented exponential growth, substantial new evidence may have emerged since. Fifth, the inclusion of reviews alongside primary studies introduces a risk of overlapping evidence, although this was deemed necessary for a comprehensive landscape analysis. Sixth, the comparative-effectiveness estimates in Figure 5 represent synthesized values across heterogeneous studies with varying methodologies, sample sizes, and outcome measures, and should be interpreted as indicative rather than definitive. Finally, publication bias whereby positive results are more likely to be published may have influenced the overall favorability of the synthesized evidence.

Conclusions

This comprehensive narrative review provides substantial evidence that AI and digital health technologies are fundamentally transforming the landscape of medical education. The convergence of LLMs, immersive technologies (VR/AR/XR), telemedicine platforms, adaptive learning systems, and simulation-based training is creating unprecedented opportunities to enhance the quality, accessibility, personalization, and efficiency of education at all levels, from undergraduate training to continuing professional development.

The evidence synthesized from 48 studies demonstrates that these technologies improve knowledge retention, clinical reasoning, procedural skills, and learner engagement relative to traditional pedagogy alone. The post-2022 exponential growth in research output, catalyzed by generative AI, reflects both the transformative potential and the urgent scholarly attention these technologies command.

This transformation is nevertheless accompanied by significant challenges that must be addressed proactively. Content accuracy and the risk of hallucination, algorithmic bias that may perpetuate health disparities, data-privacy concerns, the digital divide affecting equitable access, faculty-readiness gaps, threats to academic integrity, and the imperative to preserve the humanistic dimensions of training represent interconnected challenges that require coordinated institutional, national, and international responses.

The integrative conceptual framework proposed here offers a structured lens through which educators, institutional leaders, and policymakers can understand, navigate, and strategically leverage the AI–digital health convergence. Moving forward, responsible integration demands robust governance frameworks, rigorous longitudinal research, equity-centered implementation, comprehensive faculty development, and a sustained commitment to ensuring that technology complements rather than replaces the irreplaceable human elements of medical education and clinical care. The future of medical education lies not in wholesale adoption or rejection of these technologies, but in their thoughtful, evidence-based, and ethically guided integration into educational ecosystems that prioritize learner development, patient safety, health equity, and the cultivation of compassionate, competent healthcare professionals.

Declarations

Ethical approval and consent to participate: Not applicable

This study is a scoping review of previously published, publicly available literature and did not involve any direct participation of human subjects, human-derived material, animals, or the collection of primary data. As no human participants were enrolled and no identifiable personal data were accessed, formal ethical approval and informed consent were not required.

Consent for publication: Not applicable

This manuscript does not contain any individual person's data in any form (including individual details, images, or videos); therefore, consent for publication is not required.

Availability of supporting data

All data analyzed in this scoping review were extracted from previously published studies, which are cited in the manuscript and listed in the reference section. No new primary datasets were generated. The data charting form and the full list of studies included are available from the corresponding author upon reasonable request.

Competing interests

The author declares that he has no competing interests, financial or non-financial, that could have influenced the conduct, analysis, or reporting of this work.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. No external funding was received for the design of the study; the collection, analysis, and interpretation of data; or the writing of the manuscript.

Author Contributions

MBS - Conceptualizations, methodology, data curation, formal analysis, writing original draft, and visualization, validation, writing review and editing, and supervision. Authors reviewed and approved the final manuscript.

Use of artificial intelligence statement

The authors used an AI-assisted large language model tool to support language editing, improve readability, and assist with formatting of figures and tables. All AI-assisted output was critically reviewed, verified, and edited by the authors, who take full responsibility for the integrity, accuracy, and originality of the content.

Human ethics: Not applicable

The review synthesized data exclusively from peer-reviewed studies already in the public domain. No human participants were recruited or examined by the author, and therefore no human ethics review or institutional review board approval was sought or required for this work.

Clinical trial registration: Not applicable

This study is a scoping review of existing literature and is not a clinical trial; therefore, no clinical trial registration number is applicable.

Author contribution: MBS

Study design, Literature search, Supervision, Statistical analysis and data presentation, Data collection and manuscript Review, and finalizing the manuscript.

References

1. Gordon M, Daniel M, Ajiboye A, Hussein Uraiby, Nicole Xu Y, et al. (2024) A scoping review of artificial intelligence in medical education: BEME Guide No. 84. Med Teach 46: 446-470.
2. Ogundiya O, Rahman TJ, Valnarov-Boulter I, Young TM (2024) Looking back on digital medical education over the last 25 years and looking to the future: Narrative review. J Med Internet Res 26: e60312.
3. Daniel M, Gordon M, Patricio M, Ahmad Hider, Cameron Pawlik, et al. (2021) An update on developments in medical education in response to the COVID-19 pandemic: A BEME scoping review. Med Teach 43: 253-271.
4. Khamees D, Peterson W, Patricio M, Teresa Pawlikowska, Carolyn Commissaris, et al. (2022) Remote learning developments in postgraduate medical education in response to the COVID-19 pandemic-a BEME systematic review. Med Teach 44: 466-485.
5. Elendu C, Amaechi DC, Okatta AU, Emmanuel Amaechi C, Tochi Elendu C, et al. (2024) The impact of simulation-based training in medical education: A review. Medicine 103: e38813.
6. McGee RG, Wark S, Mwangi F, Aaron Drovandi, Faith Alele, et al. (2024) Digital learning of clinical skills and its impact on medical students' academic performance: A systematic review. BMC Med Educ 24: 1477.
7. Frenk J, Chen LC, Chandran L, Elizabeth Groff OH, Roderick King, et al. (2022) Challenges and opportunities for educating health professionals after the COVID-19 pandemic. Lancet 400: 1539-1556.
8. Shah S, Diwan S, Kohan L, David Rosenblum, Christopher Gharibo, et al. (2020) The technological impact of COVID-19 on the future of education and health care delivery. Pain Physician 23: S367-S380.
9. Sun L, Yin C, Xu Q, Zhao W (2023) Artificial intelligence for healthcare and medical education: A systematic review. Am J Transl Res 15: 4820-4828.
10. Tozsin A, Ucmak H, Soyturk S, Abdullatif Aydin, Ali Serdar Gozen, et al. (2024) The role of artificial intelligence in medical education: A systematic review. Surg Innov 31: 415-423.
11. Chan KS, Zary N (2019) Applications and challenges of implementing artificial intelligence in medical education: Integrative review. JMIR Med Educ 5: e13930.
12. Lucas HC, Upperman JS, Robinson JR (2024) A systematic review of large language models and their implications in medical education. Med Educ 58: 1276-1285.
13. Zhui L, Yhap N, Liping L, Wang Zhengjie, Xiong Zhonghao, et al. (2024) Impact of large language models on medical education and teaching adaptations. JMIR Med Inform 12: e55933.
14. Tene T, Vique López DF, Valverde Aguirre PE, Luz María Orna Puente, Cristian Vacacela Gomez (2024) Virtual reality and augmented reality in medical education: An umbrella review. Front Digit Health 6: 1365345.
15. García-Robles P, Cortés-Pérez I, Nieto-Escámez FA, Héctor García-López, Esteban Obrero-Gaitán, et al. (2024) Immersive virtual reality and augmented reality in anatomy education: A systematic review and meta-analysis. Anat Sci Educ 17: 514-528.
16. Marsilio M, Calcaterra V, Infante G, Pisarra M, Zuccotti G (2024) The digital readiness of future physicians: Nurturing the post-pandemic medical education. BMC Health Serv Res 24: 885.
17. Mainz A, Nitsche J, Weirauch V, Meister S (2024) Measuring the digital competence of health professionals: Scoping review. JMIR Med Educ 10: e55737.
18. Martinengo L, Ng MSP, Ng TDR, Yi-Ian Ang, Ahmad Ishqi Jabir, et al. (2024) Spaced digital education for health professionals: Systematic review and meta-analysis. J Med Internet Res 26: e57760.
19. Laupichler MC, Rother JF, Grunwald Kadow IC, Seifollah Ahmadi, Tobias Raupach (2024) Large language models in medical education: Comparing ChatGPT- to human-generated exam questions. Acad Med 99: 508-512.
20. Liu M, Okuhara T, Chang X, Ritsuko Shirabe, Yuriko Nishiie, et al. (2024) Performance of ChatGPT across different versions in medical licensing examinations worldwide: Systematic review and meta-analysis. J Med Internet Res 26: e60807.
21. Chowdhury PN, Vaish A, Puri B, Vaishya R (2024) Medical education technology: Past, present and future. Indian J Orthop 58: 987-999.
22. Abd-Alrazaq A, AlSaad R, Alhuwail D, Arfan Ahmed, Padraig Mark Healy, et al. (2023) Large language models in medical education: Opportunities, challenges, and future directions. JMIR Med Educ 9: e48291.
23. Siddiqui MF, Jabeen S, Alwazzan A, Sebastiano Vacca, Lara Dalal, et al. (2025) Integration of augmented reality, virtual reality, and extended reality in healthcare and medical education: A systematic review. J Med Educ Curric Dev 12: 23821205251342315.
24. Mergen M, Graf N (2024) Reviewing the current state of virtual reality integration in medical education-a scoping review. BMC Med Educ 24: 788.
25. Nagi F, Salih R, Alzubaidi M, Hurmat Shah, Tanvir Alam,

et al. (2023) Applications of artificial intelligence (AI) in medical education: A scoping review. Stud Health Technol Inform 305: 648-651.

26. Zhang W, Cai M, Lee HJ, Evans R, Zhu C, et al. (2024) AI in medical education: Global situation, effects and challenges. Educ Inf Technol 29: 4611-4633.

27. Page MJ, McKenzie JE, Bossuyt PM, Isabelle Boutron, Tammy C Hoffmann, et al. (2021) The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 372: n71.

28. Kung TH, Cheatham M, Medenilla A, Czarina Sillos, Lorie De Leon, et al. (2023) Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLoS Digit Health 2: e0000198.

29. Gilson A, Safranek CW, Huang T, Vimig Socrates, Ling Chi, et al. (2023) How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 9: e45312.

30. Takagi S, Watari T, Erabi A, Sakaguchi K (2023) Performance of GPT-3.5 and GPT-4 on the Japanese medical licensing examination: Comparison study. JMIR Med Educ 9: e48002.

31. Jin HK, Lee HE, Kim E (2024) Performance of ChatGPT-3.5 and GPT-4 in national licensing examinations for medicine, pharmacy, dentistry, and nursing: A systematic review and meta-analysis. BMC Med Educ 24: 1013.

32. Bhayana R, Krishna S, Bleakney RR (2023) Performance of ChatGPT on a radiology board-style examination. Radiology 307: e230582.

33. Mirza FN, Lim RK, Yumeen S, Samer Wahood, Bashar Zaidat, et al. (2024) Performance of three large language models on dermatology board examinations. J Investig Dermatol 144: 398-400.

34. Longwell JB, Hirsch I, Binder F, Galileo Arturo Gonzalez Conchas, Daniel Mau, et al. (2024) Performance of large language models on medical oncology examination questions. JAMA Netw Open 7: e2417641.

35. Dhanvijay AKD, Pinjar MJ, Dhokane N, Smita R Sorte, Amita Kumari, et al. (2023) Performance of large language models (ChatGPT, Bing Search, and Google Bard) in solving case vignettes in physiology. Cureus 15: e42972.

36. Cheung BHH, Lau GKK, Wong GTC, Elaine Yuen Phin Lee, Dhananjay Kulkarni, et al. (2023) ChatGPT versus human in generating medical graduate exam questions-a multinational prospective study. PLoS One 18: e0290691.

37. Kıyak YS, Emekli E (2024) ChatGPT prompts for generating multiple-choice questions in medical education and evidence on their validity: A literature review. Postgrad Med J 100: 858-865.

38. Tsang R (2023) Practical applications of ChatGPT in undergraduate medical education. J Med Educ Curric Dev 10: 23821205231178049.

39. Preiksaitis C, Rose C (2023) Opportunities, challenges, and future directions of generative artificial intelligence in medical education: Scoping review. JMIR Med Educ 9: e48785.

40. Gurnani B, Kaur K (2024) Leveraging ChatGPT for ophthalmic education: A critical appraisal. Eur J Ophthalmol 34: 323-327.

41. Sallam M (2023) The utility of ChatGPT as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations. Healthcare 11: 887.

42. Safranek CW, Sidamon-Eristoff AE, Gilson A, Chartash D (2023) The role of large language models in medical education: applications and implications. JMIR Med Educ 9: e50945.

43. Bair H, Norden J (2023) Large language models and their implications on medical education. Acad Med 98: 869-870.

44. Clusmann J, Kolbinger FR, Muti HS, Zunamys Carrero I, Jan-Niklas Eckardt, et al. (2023) The future landscape of large language models in medicine. Commun Med 3: 141.

45. Kasneci E, Sessler K, Küchemann S, Maria Bannert, Daryna Dementieva, et al. (2023) ChatGPT for good? On opportunities and challenges of large language models for education. Learn Individ Differ 103: 102274.

46. Breeding T, Martinez B, Patel H, Hazem Nasef, Hasan Arif, et al. (2024) The utilization of ChatGPT in reshaping future medical education and learning perspectives: A curse or a blessing? Am Surg 90: 560-566.

47. Cross J, Robinson R, Devaraju S, Andrea Vaughans, Ricardo Hood, et al. (2023) Transforming medical education: Assessing the integration of ChatGPT into faculty workflows at a Caribbean medical school. Cureus 15: e47014.

48. Winkler-Schwartz A, Bissonnette V, Mirchi N, Nirros Ponnudurai, Recai Yilmaz, et al. (2019) Artificial intelligence in medical education: Best practices using machine learning to assess surgical expertise in virtual reality simulation. J Surg Educ 76: 1681-1690.

49. Guerrero DT, Asaad M, Rajesh A, Abbas Hassan, Charles E Butler (2023) Advancing surgical education: The use of artificial intelligence in surgical training. Am Surg 89: 49-54.

50. González-Calatayud V, Prendes-Espinosa P, Roig-Vila R (2021) Artificial intelligence for student assessment: A systematic review. Appl Sci 11: 5467.

51. Sadek O, Baldwin F, Gray R, Allison Guerin, Hayley A Gans, et al. (2023) Impact of virtual and augmented reality on quality of medical education during the COVID-19 pandemic: A systematic review. J Grad Med Educ 15: 328-338.

52. Barteit S, Lanfermann L, Bärnighausen T, Florian Neuhann, Claudia Beiersmann, et al. (2021) Augmented, mixed, and virtual reality-based head-mounted devices for medical education: Systematic review. JMIR Serious Games 9: e29080.

53. Lewis KO, Popov V, Fatima SS (2024) From static web to metaverse: Reinventing medical education in the post-pandemic era. Ann Med 56: 2305694.

54. Vitaliy Popov, Natalie Mateju, Caris Jeske, Kadriye Lewis O (2024) Metaverse-based simulation: A scoping review of charting medical education over the last two decades. Ann Med 56: 2424450.

55. Bajra R, Frazier W, Graves L, Katherine Jacobson, Andres Rodriguez, et al. (2023) Feasibility and acceptability of a US national telemedicine curriculum for medical students and residents. JMIR Med Educ 9: e43190.

56. Lesselroth B, Monkman H, Palmer R, Craig Kuziemsky, Andrew Liew, et al. (2024) Assessing telemedicine competencies: Developing and validating learner measures for simulation-based telemedicine training. AMIA Annu Symp Proc 2023: 474-483.

57. Sharma S, Rawal R, Shah D (2023) Addressing the

challenges of AI-based telemedicine: Best practices and lessons learned. J Educ Health Promot 12: 338.

58. Wang S, Wang F, Zhu Z, Jingxuan Wang, Tam Tran, et al. (2024) Artificial intelligence in education: A systematic literature review. Expert Syst Appl 252: 124167.

59. Omiye JA, Lester JC, Spichak S, Rotemberg V, Daneshjou R (2023) Large language models propagate race-based medicine. NPJ Digit Med 6: 195.

60. Yan L, Sha L, Zhao L, Yuheng Li, Roberto Martinez-Maldonado, et al. (2024) Practical and ethical challenges of large language models in education: A systematic scoping review. Br J Educ Technol 55: 90-112.