MedPharmRes (MPR) TITLE PAGE

Upload this completed form to website with submission

ARTICLE INFORMATION	Fill in information in each box below
Article Type	Trends and Applications of AI in Competency-Based Education in Medical Programs: A Scoping Review
Article Title (within 20 words without abbreviations)	Trends and Applications of Artificial Intelligence in Competency-Based Education in Medical Programs: A Scoping Review
Running Title (within 10 words)	AI and Competency-Based Medical Programs
Author	Sang Thanh Do [1*], Cuong Ly To [1], Quoc Kha Vo Huynh [1], Duy Thanh Huynh [1], Song Thi-Thu Nguyen [1], Phuong Thi – Lan Le[1]
Affiliation	[1] Faculty of Traditional Medicine, University of Medicine and Pharmacy at Ho Chi Minh City, Vietnam
ORCID (for more information, please visit https://orcid.org)	Sang Thanh Do (https://orcid.org/0009-0009-8176-2501) Cuong Ly To (https://orcid.org/0009-0009-6925-7019) Quoc Kha Vo Huynh (https://orcid.org/0000-0002-0372-0076) Duy Thanh Huynh (https://orcid.org/ 0009-0003-2486-4356) Lan-Phuong Thi Le (https://orcid.org/0009-0008-3761-140X) Song Thi-Thu Nguyen (https://orcid.org/0009-0006-7869-7241)
Competing interests	No potential conflict of interest relevant to this article was reported.
Funding sources State funding sources (grants, funding sources, equipment, and supplies). Include name and number of grant if available.	University of Medicine and Pharmacy at Ho Chi Minh City.
Acknowledgements	This study was provided in parts by research grants from University of Medicine and Pharmacy at Ho Chi Minh city.
Availability of data and material	Upon reasonable request, the datasets of this study can be available from the corresponding author.
Authors' contributions Please specify the authors' role using this form. Authors can't change and add items, but you can delete items that are not applicable.	Conceptualization: Sang Thanh Do, Cuong Ly To, Quoc Kha Vo Huynh, Duy Thanh Huynh Data curation: Sang Thanh Do, Cuong Ly To, Quoc Kha Vo Huynh, Duy Thanh Huynh, Phuong Thi – Lan Le Formal analysis: Sang Thanh Do, Cuong Ly To, Song Thi-Thu Nguyen, Phuong Thi – Lan Le Methodology: Sang Thanh Do, Cuong Ly To, Quoc Kha Vo Huynh, Duy Thanh Huynh Validation: Sang Thanh Do, Cuong Ly To, Quoc Kha Vo Huynh, Duy Thanh Huynh Writing - original draft: Sang Thanh Do, Cuong Ly To, Duy Thanh Huynh, Phuong Thi – Lan Le Writing - review & editing: Sang Thanh Do, Cuong Ly To, Quoc Kha Vo Huynh, Duy Thanh Huynh, Song Thi-Thu Nguyen, Phuong Thi – Lan Le (All authors should be listed in this field.).
Ethics approval and consent to participate	This scoping review did not involve any direct interaction with human participants, and ethical approval was not required. However, all included studies were verified for ethical sound by adhering to the ethical guidelines provided by the journals in which they were published

CORRESPONDING AUTHOR CONTACT INFORMATION

For the corresponding author (responsible for correspondence, proofreading, and reprints)	Fill in information in each box below
First name, middle initial, last name	Sang, Thanh, Do
Email address – this is where your proofs will be sent	dtsang@ump.edu.vn
Secondary Email address	Thanhsang.do.01@gmail.com
Address	Faculty of Traditional Medicine, University of Medicine and Pharmacy at Ho Chi Minh City, 217 Hong Bang Street, District 5, Vietnam
Cell phone number	Tel: (+84) 0948561030.
Office phone number	
Fax number	

ABSTRACT

Introduction: The integration of Artificial Intelligence (AI) in medical education has emerged as a transformative shift, particularly within Competency-Based Medical Education (CBME). AI technologies, including Natural Language Processing (NLP) and machine learning, offer opportunities to enhance personalized learning and competency assessment.

Methods: A scoping review was conducted following the framework by Arksey and O'Malley (2005) to examine the current integration of AI in CBME. Empirical studies were included, focusing on AI applications in medical education, competency assessments, and skill development.

Results: The 50 studies, published from 2010 to 2025, were included in the scoping review and the synthesized evidence demonstrated that AI has shown potential in automating assessments, providing real-time feedback, and supporting personalized learning paths. Common AI technologies such as generative AI, NLP, and machine learning were applied across diverse medical education settings. However, challenges regarding ethical concerns, faculty training, and limited integration within established curricula were identified.

Conclusion: The integration of AI into CBME offers significant potential in medical education; however, several challenges remain. There is a need for more empirical research, longitudinal studies, and AI literacy programs such as training in prompt engineering, AI ethics, and responsible data use for both educators and students. Addressing these gaps will ensure AI's effective, ethical, and equitable integration in medical training.

Keywords: Artificial Intelligence, Competency-Based Education, Teaching

1. Introduction

The integration of Artificial Intelligence (AI) in healthcare education represents a transformative shift, responding to the demands of a rapidly evolving healthcare landscape and the growing complexity of medical practice [1]. Historically, medical education has focused on providing knowledge and clinical skills through time-based curricula. However, with the increasing imperative to optimize service delivery and address disparities in medical training, there is a shifting focus towards Competency-Based Education, which centers on learners' ability to demonstrate specific competencies rather than simply completing a set number of hours or procedures [2]. This educational transformation seeks to ensure that medical students are adequately prepared for authentic clinical practice environments by focusing on proficiency in core competencies, ethical conduct and professional behavior, and other vital clinical attributes at each stage of the educational continuum [3-5].

With the rapid advancement of AI, novel opportunities have emerged to strengthen the Competency-Based Medical Education (CBME) model. AI systems can facilitate the formative and summative assessment of student competencies through real-time feedback, enabling personalized learning trajectories, and utilizing advanced analytics to monitor learners' progress across multiple domains of professional development [6, 7].

AI systems, including those powered by Natural Language Processing (NLP) and machine learning, automate labor-intensive tasks such as grading, supporting the evaluation in the assessment of clinical skills, and even predicting individual learning gaps for students based on longitudinal performance data [8, 9]. This could significantly alleviate the workload for medical educators while promoting greater objectivity and consistency in evaluations [10].

The integration of AI in CBME represents a promising development in medical education, offering the ability to assess both cognitive and non-cognitive aspects of learning. AI's capacity to harness large-scale data analytics can generate nuances into student performance, identify areas in need of remediation, and support the creation of adaptive, personalized learning trajectories [11-13]. For example, AI-enhanced simulation platforms (such as MedSimAI) are capable of replicating authentic clinical encounters and delivering immediate, structured feedback aligned with standardized assessment instruments, such as the Master Interview Rating Scale (MIRS). These systems facilitate the evaluation of both cognitive competencies (e.g., clinical reasoning) and non-cognitive attributes (e.g., communication skills and empathy) [11].

Despite evident advantages of AI integration in medical education, the implementation is accompanied by several challenges. Key concerns include the ethical implications of employing AI in educational assessments, the reliability and validity of AI-generated feedback, and the need for comprehensive training for both educators and learners to effectively use AI tools. Furthermore, medical institutions often encounter systemic barriers, such as limited financial and infrastructural resources, insufficient faculty expertise on AI technologies, and institutional resistance to pedagogical change [14].

In response to the emerging opportunities and ongoing challenges, a scoping review is needed to investigate the current state of AI integration within CBME. In particular, the study must explore the AI technologies being utilized, their applications in competency assessment and skill development, and the extent of their impact on medical education [1]. Therefore, this scoping review aimed to analyze published literature on AI in CBME to identify current application trends, categorize prevalent AI technologies, and highlight existing research gaps requiring further exploration. By systematically analyzing the existing body of literature, this review mapped prevailing practices, identified research gaps, and outlined future directions to enhance the effective and ethical implementation of AI in medical education. Furthermore, this review seeks to elucidate the role of AI in supporting competency evaluation within CBME frameworks by examining its applications in the assessment of cognitive functions, procedural proficiency, and interpersonal communication skills.

2. Materials and Methods

2.1. Review Design and Framework.

This scoping review adopted the methodological framework developed by Arksey and O'Malley (2005) [15] and refined by Levac et al. (2010) [16], and adhered to the PRISMA extension for Scoping Reviews (PRISMA-ScR) Checklist (Appendix 1). The primary objectives were to map the scope of AI applications in CBME across medical training programs, and to examine the extent to which AI technologies are being integrated into assessment processes and competency development. Publications included in the review were empirical studies on the use of AI tools in both formative and summative evaluations within CBME frameworks, offering a comprehensive overview of AI-enabled approaches to assess clinical performance, cognitive abilities, and other essential competencies in medical education.

2.2. Inclusion and Exclusion Criteria

The studies included in this review were selected according to the criteria presented in Table 1.

Insert Table 1

2.3. Information Sources and Search Strategy

The literature search was conducted on five databases (i.e. PubMed, Scopus, ERIC, ScienceDirect, and Google Scholar) for peer-reviewed English-language publications up to April 2025. The search strategy was formed by keywords across three domains:

(1) Artificial Intelligence
(2) Competency-Based Education
(3) Medical Education

The search query was developed by combining the keywords using Boolean operators as follows:

("artificial intelligence" OR "machine learning" OR "deep learning" OR "natural language processing" OR "generative AI" OR "ChatGPT") AND ("competency-based education" OR "CBME") AND ("medical education" OR "medical training" OR "clinical education")

Truncation (e.g., educat*) and filters were applied when supported by the database. Reference lists of included articles were also manually screened to capture additional relevant studies.

2.4. Study Selection

After initial screening, studies were assessed based on titles and abstracts using predefined inclusion criteria. Full-text articles were then reviewed by two independent reviewers. To enhance objectivity, the Rayyan platform was used to facilitate blinded screening, preventing bias between reviewers [17]. Any discrepancies were discussed and resolved by a third senior reviewer with expertise in medical education and AI, ensuring consistent judgment. Although interrater agreement statistics were not calculated, consensus was achieved on all included studies.

2.5. Data Extraction

Data were extracted from the included studies using a standardized data extraction form, which captured the following information: Study characteristics (e.g., authors, year of publication, study design); AI applications in medical education (e.g., natural language processing (NLP), machine learning, AI-driven feedback systems); Competency domains assessed (e.g., medical knowledge, clinical skills, interpersonal communication, professionalism); Educational settings (e.g., undergraduate, postgraduate, clinical training, virtual learning environments); Impact of AI on the competency development (e.g., improved feedback, personalized learning paths, assessment efficiency); Challenges and barriers to AI adoption in CBME (e.g., ethical considerations, faculty training, data privacy).

2.6. Data Analysis and Synthesis

To explore trends over time, a year-wise descriptive analysis was conducted on the included studies. For each publication year, the following variables were aggregated: the number of studies, countries involved, research designs, types of AI, and areas of application in medical and health professions education. This data set was synthesized into a structured table, providing a comprehensive longitudinal overview of how AI technologies have been integrated into medical education across geographic, methodological, and thematic dimensions. Data cleaning included normalization of country names, grouping of similar AI techniques (e.g., Diffusion Models, Large Language Models), and clustering study design terms for consistency. To visualize the global distribution of studies, a choropleth mapping analysis was performed in this review. Country information was extracted from each article and standardized using codes for representation of country names by ISO. The number of studies originating from or conducted in each country was aggregated. The results were plotted on a world map using a color gradient to indicate frequency, with darker colors representing higher research activity. This visualization was implemented in R using the 'rworldmap' and 'dplyr' packages**.** The data analysis was conducted using a descriptive summary and thematic synthesis approach to accommodate the diversity of study designs and outcomes included in this research. Quantitative data were summarized using appropriate descriptive statistics such as means, standard deviations, frequencies, and percentages to illustrate the distribution and characteristics of the sample. Qualitative data were analyzed thematically, identifying recurrent themes and patterns relevant to the research objectives. To ensure rigor, data coding was performed independently by two researchers, and discrepancies were resolved through discussion to reach a consensus. The synthesis process involved grouping the findings into thematic categories that reflect the core aspects of the study, allowing for a comprehensive understanding of the phenomena under investigation. The main themes included: [Theme 1: e.g., types of interventions or technologies], [Theme 2: e.g., impacts or outcomes observed], [Theme 3: e.g., challenges or barriers], and [Theme 4: e.g., recommendations and future directions]. Each theme was explored in detail to highlight similarities and variations across the studies. Where applicable, subgroup analyses or comparative assessments were conducted to examine differences by corresponding variables such as demographic factors, study settings, or intervention types. The integration of quantitative and qualitative findings enabled a holistic interpretation, supporting the development of evidence-based conclusions. All statistical analyses and data visualization were performed using the R environment (version 4.5.1 in 2025), ensuring reproducibility and robust data handling, with results presented in tables and figures to facilitate clarity and comprehension. A narrative summary was then compiled, providing insights into the strengths and weaknesses of current AI applications in medical education.

2.7. Quality Assessment

Given the scope of the review and the heterogeneity of included studies, a formal quality assessment of individual studies was not conducted. However, the methodological rigor of each study was noted in the data extraction process. The procedure of this scoping review was prospectively registered on the Open Science Framework (OSF) under the registration DOI: https://doi.org/10.17605/OSF.IO/9UPHE. The protocol followed the methodological framework proposed by Arksey and O'Malley (2005) [15] and refined by Levac et al. (2010) [16], ensuring transparency and reproducibility of the review process.

2.8. Ethical Considerations

This scoping review did not involve any direct interaction with human participants, and ethical approval was not required. However, all included studies were verified for adherence to ethical standards according to the guidelines of the corresponding journals (in which they were published).

3. Results

3.1. Summary of Scoping Review Findings

Insert Figure 1

Insert Table 2

A total of 50 studies, published between 2010 and 2025, were included in the review, extracted from five major databases: ERIC (n = 64), PubMed (n = 12), Google Scholar (n = 64), ScienceDirect (n = 67), and Scopus (n = 32), following the PRISMA flowchart (Figure 1). Regarding the year-wise analysis, there has been a rapid increase in publications since 2019, with a peak of 25 studies in 2024.

Early studies (2010–2020) mainly originated from the USA, UK, and Canada, employing machine learning and deep learning techniques, often through quantitative or review-based designs, focused on foundational curriculum assessment and surgical video evaluation. Between 2021 and 2023, research distribution expanded geographically to China, Malaysia, Pakistan, Saudi Arabia, and several European countries. During this period, AI types diversified to include intelligent tutoring systems, learning analytics, and generative AI/large language models (GenAI/LLMs). Study designs encompassed qualitative study, experimental empirical study, systematic reviews, and perspective papers. Application areas broadened to nursing education, pharmaceutical education, and public health contexts. By 2024, GenAI and LLMs became dominant, with multiple studies reporting applications in nursing, digital health education, surgical training, and pediatrics. The USA remained the leading contributor, followed by France, Germany, India, and Australia. The most recent studies up to April 2025 included emerging publications from the United Arab Emirates, Spain, Ghana, and the UK, focusing on clinical competency assessments, radiology, and endoscopy education (Table 2).

3.2. Geographic Trends in AI Applications in Medical Education

Insert Figure 2

Insert Figure 3

The world map highlights the geographic distribution of research on AI integration in medical education. The USA accounted for the highest number of studies, followed by Canada, the UK, Germany, India, and China. Several countries in Europe, Asia, and the Middle East also contributed to this body of research, although with fewer publications. The map demonstrates a clear dominance of high-income countries in the technological development and publication of AI-related educational innovations. In contrast, regions such as Africa and parts of South America remain underrepresented in the current literature (Figure 2). In addition, most studies originated from the USA, followed by Canada, Germany, Taiwan, and the UK (Figure 3.a). Narrative and conceptual reviews were the most common study designs (Figure 3.b). Publications increased sharply in 2024 (Figure 3.c), reflecting rising interest in AI integration for CBME. The top publishing journals included *JMIR Medical Education*, *BMC Medical Education*, and *Academic Medicine*, highlighting both educational and interdisciplinary relevance (Figure 3.d).

3.3. Thematic and Temporal Trends of AI Applications in Medical Education

Insert Table 3

The analysis identified five major topics in AI applications in medical education. These topics are summarized in Table 3, along with their associated top keywords and applications: **Topic 1** (Focused on *ChatGPT, machine learning, deep learning*, and *LLMs*): The key applications included Medical Education in Ophthalmology, Training Medical Students in AI Concepts, and CBME. **Topic 2** (Emphasized *GenAI, machine learning*, and *support systems*): This topic had applications in clinical decision support, AI-enhanced curriculum design, and assessment frameworks. **Topic 3** (Highlighted *natural language processing (NLP), LLMs*, and *learning analytics*): These were predominantly applied in personalized learning systems, feedback analysis, and interactive learning environments. **Topic 4** (Centered on *deep learning, GenAI*, and *artificial intelligence systems*): The applications included simulation-based education, intelligent tutoring systems, and virtual healthcare environments. **Topic 5** (Focused on *NLP, large language models*, and *chatbot technologies*): Applications were found in automated essay-scoring, diagnostic decision-making, and patient feedback systems.

Insert Figure 4

The temporal analysis (Figure 4) demonstrates a significant increase in AI-related publications in medical education, particularly since 2020, with a peak observed in 2024. Topic 5 (centered on LLMs, NLP, and ChatGPT) has dominated recent years, especially from 2023 to 2024. Topic 1, though less frequent overall, appeared to be concentrated in 2024. Geographically, the USA led across all topics, notably in Topic 1 and Topic 5. Countries such as Germany, India, France, Canada, and the UK also contributed across multiple themes. This distribution highlights both the global engagement in AI-enhanced medical education and the growing concentration of research efforts on GenAI and language-based tools in recent years.

3.4. Mapping AI Technologies to Educational Applications: A Sankey Visualization

Insert Figure 5

The Sankey diagram (Figure 5) represents the predominant artificial intelligence methodologies utilized across diverse domains of medical education. Machine learning, deep learning, and GenAI/LLMs constituted the most prevalent approaches, exhibiting strong associations with general medical education, surgical training, and diagnostic instruction. Less commonly applied, but nonetheless, noteworthy AI techniques included NLP, intelligent tutoring systems, and ontology-based rule-driven frameworks, which are primarily linked to specialized fields such as pediatric education, pharmaceutical education, and residency training programs. This visualization highlights the breadth and heterogeneous adoption patterns of AI applications across educational domains, emphasizing a nascent approach in well-established research and pedagogical innovation.

3.5. Literature Gaps in AI Integration within CBME

Insert Figure 6

The word cloud generated from the "Gap Identified" field in our scoping review offers a concise and visual-intuitive overview of the most-frequently reported limitations and unmet needs across the included studies (Figure 6). Several key research gaps emerged as dominant themes. First, there is a lack of standardization, particularly in methodology, evaluation metrics, and integration frameworks for AI tools in medical education. Second, ethical concerns related to data privacy, informed consent, and responsible deployment of AI were frequently highlighted. Third, technical limitations were noted, including gaps in language processing capabilities, ontology modeling, and adaptive personalization of AI systems. Fourth, many studies emphasized the need for summative evaluation, scalable implementation strategies, and real-time decision support, especially in tool development and deployment phases. Fifth, a shortage in full-scale clinical validation and real-world testing of AI-integrated educational interventions was commonly reported.

4. Discussion

This scoping synthesis illustrates a dynamic and maturing research landscape in the application of AI within medical education, particularly within the emerging CBME approach that has not been previously implemented. The review is characterized by a sharp rise in publication volume, diversification of AI applications, and broader geographical engagement. Since 2019, there has been an exponential increase in scholarly output, peaking in 2024. Overall, this trend reflects a maturing research landscape by increasing international collaboration, methodological diversity, and expanding AI applications in medical education, which involve both technological advances and the rising urgency to modernize medical curricula in an era of digital transformation. The prevailing research designs contributed theoretical values, including narrative reviews, perspective papers, and cross-sectional surveys. This indicates that the field is currently in an "exploratory phase", with only a limited number of experimental studies featuring control groups or randomized controlled trials (RCTs). The most commonly leveraged technologies are GenAI and LLMs, notably ChatGPT, along with NLP, traditional machine learning, and learning analytics. These technologies have been applied across the entire medical education pipeline, including clinical simulation, personalized learning pathways, qualitative feedback analysis, and CBME assessments through EPAs. The most significant gap in the current literature lies within the absence of randomized controlled trials, the lack of long-term followup studies, and the limited evidence on cost-effectiveness outcomes. These findings are consistent with the trend highlighted by Domrös-Zoungrana et al. (2024) [63], who emphasized that most current evidence remains descriptive, with limited availability of empirical data. A novel observation is the rapid shift toward GenAI and LLMs during 2023–2024, which has clearly surpassed the prevalence of traditional NLP reported in reviews published prior to 2022. Moreover, geographic mapping reveals a stark imbalance: regions such as Africa and South America are largely absent from the research landscape. This omission has not been adequately addressed in previous reviews.

4.1. Evolution and Maturity of Research

Early studies (2010–2020) predominantly emerged from North America and Western Europe, focusing on foundational machine learning applications in curriculum evaluation [18, 19] and video analysis for surgical training [20]. These efforts laid the groundwork for methodological rigor and established AI's potential in educational settings. Post-2020 research rapidly expanded to Asia [25, 36, 59], the Middle East [32, 42], and Africa [60, 62], signifying a global recognition of AI's relevance and a democratization of research outputs. Countries like China [37, 51, 59], India [31, 46, 48, 65], Saudi Arabia [29, 32], and Malaysia [29] now feature prominently in the literature.

4.2. Shifts in AI Modalities and Study Designs

The field has evolved from a reliance on traditional machine learning and deep learning to the adoption of GenAI and LLMs [31, 37, 40-42, 48, 49], such as ChatGPT [61-63]. These newer tools have demonstrated potential across diverse educational applications, ranging from automated essay grading to OSCE evaluations [25, 61]. Additionally, the spectrum of research designs has broadened, encompassing qualitative analyses, mixed-method studies, and systematic reviews, reflecting the growth in methodological sophistication and contextual adaptation.

4.3. Geographic and Topic Distribution

A choropleth and Sankey analysis underscore the dominance of high-income countries, notably the USA, Germany, Canada, and UK, across AI topics. Despite modest contributions, nations like Ghana and the UAE signaled the emergence of new research hubs. Topic modeling reveals five dominant themes, with recent years heavily weighted toward LLMs, NLP, and GenAI, especially in personalized learning, clinical simulation, and decision support systems.

4.4. From the Landscape to Practical Implications

The topic modeling in this study revealed five core themes reflecting distinct applications of AI in medical education. This pattern is consistent with recent literature.

Topic 1 on the integration of LLMs and ChatGPT into CBME and EPA framework as demonstrated in the study by Roy, Asitava Deb et al. (2024) [46], where ChatGPT generated appropriate responses in AETCOM scenarios, and supported both learners and educators in resource-limited settings. This aligns with findings by Kung et al. (2023), which showed that ChatGPT achieved passing scores on USMLE-style questions, highlighting its potential for self-directed clinical reasoning [66]. Sallam (2023) also reported its educational potential in reflective learning and case-based thinking [67]. A multinational study by Wang et al. (2023), involving authors from Vietnam, China, Cambodia, Singapore, US, Austria, Poland, and UK, focused on low and middle-income country setting [37]. The study highlighted ChatGPT's potential to enhance access to medical information, improve public health literacy, support telemedicine, and translate medical content into local languages. This technology holds promises for advancing global health equity, facilitating remote medical education, and improving care in underserved regions. Its integration into emerging models such as CBME may help overcome language and resource barriers in settings like Southeast Asia.

In **Topics 2–4**, which emphasize the use of machine learning and GenAI for simulation and feedback, are supported by findings from Jason A. Reid et al. (2025). Their study demonstrated that ChatGPT-3 enabled nursing educators to rapidly generate complex clinical scenarios, promote critical thinking, and significantly reduce simulation-development time. While ChatGPT proved to be a creative, convenient, and efficient tool for nursing education, the study also highlighted the need for ongoing evaluation of the medical accuracy and currency of AI-generated content [38]. Pathiyil Ravi Shankar et al. (2023) reported that medical faculty from Malaysia and Pakistan viewed time-flexible and competency-personalized undergraduate medical programs as feasible, enabling students to progress at their own pace upon demonstrating required competencies. AI and online learning were regarded as essential tools to support this individualized educational approach [38]. Additionally, Alison Lentz et al (2021) introduced the concept of "AI-assessment", which integrates artificial intelligence into medical education assessment systems to shift from summative evaluation toward formative and learning-oriented feedback. AI was shown to facilitate continuous feedback, personalize learning trajectories, and support reliable assessment of professional activities (EPAs). However, the authors emphasized that AI is not inherently objective and should be co-developed with stakeholders to ensure fairness and educational effectiveness [24].

Topic 5 focuses on NLP and chatbots as virtual assistants aligns with the review by Booth et al. (2022), who developed a natural language processing model to automatically classify faculty feedback into sub-competencies based on the Milestone 2.0 framework in medical education. Trained across three institutions and validated at a fourth, the model demonstrated high accuracy in domains such as professionalism, communication, and lifelong learning. It enabled the rapid organization of hundreds of feedback entries and generated self-assessment reports within one minute, and was successfully deployed in a functional web-based application [28]. These parallels support our interpretation that AI tools are no longer experimental novelties, but evolving educational supports structured around key pedagogical goals.

4.6. Temporal and Geographic Trends

The temporal analysis in the study shows a sharp rise in AI-related publications between 2020 and 2024, which aligns with the COVID-19 pandemic's push toward digital transformation in education [68]. The United States takes a lead in AI-integration research for medical education, a trend consistent with findings reported by Shuang Wang et al. in 2024 [69], followed by Germany, India, France, Canada, and the UK. This pattern likely reflects disparities in infrastructure, research funding, and policy support [42]. Notably, emerging countries like India [46] and Brazil are gaining prominence, particularly in **Topic 5** applications such as chatbots for public health education, suggesting a diversification of innovation sources. These trends highlight both the global expansion and the persistent imbalance in AI research capacity across regions.

4.7. Research Gaps and Priority Actions

Several critical gaps remain in the integration of AI into medical education, each with distinct implications for practice and policy. First, the lack of standardized AI competency frameworks for faculty and students contributes to fragmented teaching and limits alignment with CBME. This gap is well-recognized by Mitchell G. Goldenberg in 2023 [70], who called for the development of structured AI literacy curricula in undergraduate medical education. Second, the scarcity of RCTs and long-term evaluations limits our ability to determine true effectiveness or cost-benefit. This issue has also been acknowledged in previous studies by Fatima et al. in 2024 [42] and Booth et al. in 2023 [28]. Third, the integration of AI into CBME workflows remains superficial. While tools like LLMs are being explored for feedback and decision-making support, few studies demonstrate their systematic incorporation into EPA-based assessment systems, a challenge noted by Booth et al. in 2023 [24]. Although NLP was effective in mapping feedback to Milestone sub-competencies, its operational integration into momEPA dashboards or entrustment workflows was not evaluated [28]. Current AI ethics and validation protocols are often lacking, posing safety concerns. As highlighted by Xin Wang et al. in 2023, while ChatGPT supports basic medical education, concerns remain regarding content accuracy, legal liability, and ethical use in practice [71]. Addressing these gaps will be critical to ensure AI contribution in not only innovation, but also safety, equity, and educational value to competency-based training.

4.8. Implications for Policymakers and Educators

To operationalize AI integration in medical education, particularly within CBME that emphasizes personalized and learner-centered approaches, this study proposes several actionable strategies grounded in recent trends and supported by emerging research. First, developing institutional AI sandboxes can provide students and educators with controlled environments to engage in prompt engineering, clinical case simulation, and generation of standardized assessment data. Second, offering targeted micro-credentials such as *Prompt Engineering for Clinicians*, will be essential to upskill the faculty and ensure sustainable micro-credentials leadership in the digital transformation era [42]. Third, academic and industry collaboration should be fostered to develop large-scale, multilingual and ethnically diverse datasets. This approach would enhance model generalizability and reduce bias, addressing concerns raised by Mehrabi et al. (2021) regarding demographic inequities in training data[72]. To ensure both educational efficacy and ethical integrity, a two-tier evaluation framework for AI integration in medical education is proposed: 1) Learning Quality – assessed through analytics aligned with EPAs, enabling competency-based tracking of learner performance; 2) Algorithmic Governance, evaluated based on key dimensions such as model explainability, data privacy, and fairness. This dual-layered approach provides a comprehensive foundation for monitoring AI-driven educational innovations, ensuring that implementations are not only pedagogically sound but also aligned with technical transparency and ethical standards [73]. Together, these recommendations form a roadmap toward safe, equitable, and competency-driven AI deployment in medical education.

4.9. Limitations of the Scoping Review

This review has several limitations. First, the inclusion of only English-language publications may have excluded relevant non-English or locally published studies. Second, the absence of a formal quality appraisal procedure limited the certainty of the findings. Third, reliance on automated data extraction from titles and abstracts may have introduced semantic errors. Future research should address these gaps through multilingual searches, rigorous empirical studies, ethical safeguards, cost-effectiveness analyses, and the development of validated AI competency standards.

5. Conclusion

5. 1. Summary of Findings

This scoping review highlights a dynamic and increasingly sophisticated landscape of AI applications in medical education, which underscores the value of systematically mapping key concepts, evidence types, and knowledge gaps within a defined area of inquiry, marked by a growing integration of generative AI and large language models into competency-based educational frameworks. While notable progress has been made in diversifying research across regions and AI modalities, current evidence remains largely theoretical. There is limited empirical validation, a lack of long-term outcome evaluations, and an insufficient number of randomized controlled trials as well as standardized guidelines and policies regarding ethical standards. Moreover, regional disparities in research contributions and gaps in ethical oversight underscore the urgent need for inclusive, equitable, and safe implementation of AI in medical training.

5. 2. Actionable Implications and Policy Directions

To advance the responsible integration of AI into medical education, we emphasized the need for rigorous, context-sensitive research, including randomized controlled trials, cost-effectiveness studies, and the development of AI literacy frameworks such as microcredentialing in prompt engineering, AI ethics, and responsible data use. We also advocated for a dual-focus evaluation strategy that incorporates both learning outcome tracking and algorithmic governance, thereby ensuring that innovation aligns with pedagogical integrity and ethical accountability. Collectively, these insights provide a foundation for evidence-informed policymaking, responsible educational reform, and the future advancement of digital transformation in medical education.

6. Abbreviation

AI	Artificial Intelligence
GenAI	Generative Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
LLM	Large Language Model
CBME	Competency-Based Medical Education
EPA	Entrustable Professional Activity
NLP	Natural Language Processing
ITS	Intelligent Tutoring System
AETCOM	Attitude, Ethics, and Communication
MIRS	Master Interview Rating Scale
OSF	Open Science Framework
PRISMA-ScR	Preferred Reporting Items for Systematic Reviews and Meta-Analyses – Scoping Review
RCT	Randomized Controlled Trial
GPT	Generative Pre-trained Transformer
OSCE	Objective Structured Clinical Examination

REFERENCE

1. Rincón EHH, Jimenez D, Aguilar LAC, Flórez JMP, Tapia ÁER, Peñuela CLJ. Mapping the use of artificial intelligence in medical education: a scoping review. BMC Medical Education. 2025;25(1):526. doi: 10.1186/s12909-025-07089-8.
2. Ten Cate O. Competency-Based Postgraduate Medical Education: Past, Present and Future. GMS J Med Educ. 2017;34(5):Doc69. Epub 2017/12/12. doi: 10.3205/zma001146. PubMed PMID: 29226237; PubMed Central PMCID: PMCPMC5704607.
3. Bhattacharya S. Competency-based medical education: An overview. Annals of Medical Science & Research. 2023;2(3).
4. Au LN, Do MTN, Nguyen HDP. Using team-based learning for post-graduate training: challenges and solutions. MedPharmRes. 2024;8(1):1-9. doi: 10.32895/UMP.MPR.8.1.3.
5. Binh PDU, An PL, Tuan TD, Leppink J. Content validity of a professionalism OSCE developed in family medicine training University of Medicine and Pharmacy at Ho Chi Minh city Vietnam. MedPharmRes. 2018;2(1):6-11. doi: 10.32895/UMP.MPR.2.1.6.
6. Alanzi T, Alhajri A, Almulhim S, Alharbi S, Alfaifi S, Almarhoun E, et al. Artificial Intelligence and Patient Autonomy in Obesity Treatment Decisions: An Empirical Study of the Challenges. Cureus. 2023;15(11):e49725. Epub 2024/01/02. doi: 10.7759/cureus.49725. PubMed PMID: 38161816; PubMed Central PMCID: PMCPMC10757560.
7. Masters K, Heather M, Jennifer B, Tamara C, Kataryna N, Sofia V-A, et al. Artificial Intelligence in Health Professions Education assessment: AMEE Guide No. 178. Medical Teacher.1-15. doi: 10.1080/0142159X.2024.2445037.
8. Shakur AH, Holcomb MJ, Hein D, Kang S, Dalton TO, Campbell KK, et al. Large Language Models for Medical OSCE Assessment: A Novel Approach to Transcript Analysis. arXiv preprint arXiv:241012858. 2024.
9. Gupta N, Khatri K, Malik Y, Lakhani A, Kanwal A, Aggarwal S, et al. Exploring prospects, hurdles, and road ahead for generative artificial intelligence in orthopedic education and training. BMC Medical Education. 2024;24(1):1544. doi: 10.1186/s12909-024-06592-8.
10. Sriram A, Ramachandran K, Krishnamoorthy S. Artificial Intelligence in Medical Education: Transforming Learning and Practice. Cureus. 2025;17(3).

11. Hicke Y, Geathers J, Rajashekar N, Chan C, Jack AG, Sewell J, et al. MedSimAI: simulation and formative feedback generation to enhance deliberate practice in medical education. arXiv preprint arXiv:250305793. 2025.
12. Tao G, Pan Y. AI-Driven Personalized Learning in Medical Education: Enhancing Cognitive Skills and Addressing Language Proficiency Challenges. Forum for Linguistic Studies. 2025;7(4):366-84. doi: 10.30564/fls.v7i4.8932.
13. Narayanan S, Ramakrishnan R, Durairaj E, Das A. Artificial intelligence revolutionizing the field of medical education. Cureus. 2023;15(11).
14. Mahmud A, Dwivedi G, Chow BJW. Exploring the Integration of Artificial Intelligence in Cardiovascular Medical Education: Unveiling Opportunities and Advancements. Canadian Journal of Cardiology. 2024;40(10):1946-9. PubMed PMID: rayyan-206465011.
15. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International journal of social research methodology. 2005;8(1):19-32.
16. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implementation Science. 2010;5(1):69. doi: 10.1186/1748-5908-5-69.
17. Mak S, Thomas A. Steps for conducting a scoping review. J Grad Med Educ. 2022;14(5):565-7.
18. Chen C-K. Curriculum Assessment Using Artificial Neural Network and Support Vector Machine Modeling Approaches: A Case Study. IR Applications. Volume 29. Association for Institutional Research, 2010.
19. Williams P. Does Competency-Based Education with Blockchain Signal a New Mission for Universities? Journal of Higher Education Policy and Management. 2019;41(1):104-17. doi: 10.1080/1360080X.2018.1520491. PubMed PMID: rayyan-206902181.
20. Leo J, Kurdi G, Matentzoglu N, Parsia B, Sattler U, Forge S, et al. Ontology-Based Generation of Medical, Multi-Term MCQs. International Journal of Artificial Intelligence in Education. 2019;29(2):145-88. doi: 10.1007/s40593-018-00172-w. PubMed PMID: rayyan-206902225.
21. Bilgic E, Valanci-Aroesty S, Fried GM. Video Assessment of Surgeons and Surgery. Advances in Surgery. 2020;54:205-14. doi: 10.1016/j.yasu.2020.03.002. PubMed PMID: rayyan-207231763.

22. Lajoie SP. Student Modeling for Individuals and Groups: The BioWorld and HOWARD Platforms. International Journal of Artificial Intelligence in Education. 2021;31(3):460-75. doi: 10.1007/s40593-020-00219-x. PubMed PMID: rayyan-206902218.
23. Piotrkowicz A, Wang K, Hallam J, Dimitrova V. Data-Driven Exploration of Engagement with Workplace-Based Assessment in the Clinical Skills Domain. International Journal of Artificial Intelligence in Education. 2021;31(4):1022-52. doi: 10.1007/s40593-021-00264-0. PubMed PMID: rayyan-206902220.
24. Lentz A, Siy JO, Carraccio C. AI-ssessment: Towards Assessment As a Sociotechnical System for Learning. Acad Med. 2021;96(7):S87-S8. doi: 10.1097/ACM.0000000000004104. PubMed PMID: rayyan-206913215.
25. Jiang Z, Shi D, Distefano C. A Short Note on Optimizing Cost-Generalizability via a Machine-Learning Approach. Educational and Psychological Measurement. 2021;81(6):1221-33. doi: 10.1177/0013164421992112. PubMed PMID: rayyan-206902232.
26. Warren AE, Tham E, Abeysekera J. Some Things Change, Some Things Stay the Same: Trends in Canadian Education in Paediatric Cardiology and the Cardiac Sciences. CJC Pediatric and Congenital Heart Disease. 2022;1(5):232-40. doi: 10.1016/j.cjcpc.2022.08.004. PubMed PMID: rayyan-207231736.
27. Nizhenkovska IV, Reva TD, Chkhalo OM, But IO, Manchenko OV. Best Practices for Teaching Chemistry Disciplines to Graduates Majoring in Pharmacy during the COVID-19 Restrictions: A Systematic Review. International Journal of Educational Methodology. 2022;8(4):769-81. PubMed PMID: rayyan-206902224.
28. Booth GJ, Ross B, Cronin WA, McElrath A, Cyr KL, Hodgson JA, et al. Competency-Based Assessments: Leveraging Artificial Intelligence to Predict Subcompetency Content. Academic Medicine. 2023;98(4). PubMed PMID: rayyan-206464969.
29. Shankar PR, Azhar T, Nadarajah VD, Er HM, Arooj M, Wilson IG. Faculty perceptions regarding an individually tailored, flexible length, outcomes-based curriculum for undergraduate medical students. Korean journal of medical education. 2023;35(3):235-47. doi: 10.3946/kjme.2023.262. PubMed PMID: rayyan-206464977.
30. Fazlollahi AM, Yilmaz R, Winkler-Schwartz A, Mirchi N, Ledwos N, Bakhaidar M, et al. AI in surgical curriculum design and unintended outcomes for technical competencies in

simulation training. JAMA network open. 2023;6(9):e2334658-e. PubMed PMID: rayyan-206464998.

31. Ghosh A, Bir A. Evaluating ChatGPT's ability to solve higher-order questions on the competency-based medical education curriculum in medical biochemistry. Cureus. 2023;15(4). PubMed PMID: rayyan-206465000.
32. Mir MM, Mir GM, Raina NT, Mir SM, Mir SM, Miskeen E, et al. Application of artificial intelligence in medical education: current scenario and future perspectives. Journal of advances in medical education & professionalism. 2023;11(3):133. PubMed PMID: rayyan-206465013.
33. Wolf RR, Wolf AB. Using AI to Evaluate a Competency-Based Online Writing Course in Nursing. Online Learning. 2023;27(3):41-69. PubMed PMID: rayyan-206902182.
34. Pizzolla I, Aro R, Duez P, De Lièvre B, Briganti G. Integrating Artificial Intelligence into Medical Education: Lessons Learned from a Belgian Initiative. Journal of Interactive Learning Research. 2023;34(2):401-24. PubMed PMID: rayyan-206902200.
35. Rozario SY, Sarkar M, Farlie MK, Lazarus MD. Responding to the healthcare workforce shortage: A scoping review exploring anatomical pathologists' professional identities over time. Anatomical Sciences Education. 2024;17(2):351-65.
36. Srikong M, Wannapiroon P, Nilsook P. Crystallized Intelligence Wisdom Repository Management System with a Conversational Agent. International Education Studies. 2023;16(2):150-63. PubMed PMID: rayyan-206902235.
37. Wang X, Sanders HM, Liu Y, Seang K, Tran BX, Atanasov AG, et al. ChatGPT: promise and challenges for deployment in low- and middle-income countries. The Lancet Regional Health - Western Pacific. 2023;41:100905. doi: 10.1016/j.lanwpc.2023.100905. PubMed PMID: rayyan-207231766.
38. Reid JA. Building Clinical Simulations With ChatGPT in Nursing Education. The Journal of nursing education. 2024:1-2. doi: 10.3928/01484834-20240424-05. PubMed PMID: rayyan-206464976.
39. Grosjean J, Dufour F, Benis A, Januel J-M, Staccini P, Darmoni SJ. Digital Health Education for the Future: The SaNuRN (Santé Numérique Rouen-Nice) Consortium's Journey. JMIR Medical Education. 2024;10(10):e53997-e. doi: 10.2196/53997. PubMed PMID: rayyan-206464978.

40. Campbell KK, Holcomb MJ, Vedovato S, Young L, Danuser G, Dalton TO, et al. Applying State-of-the-Art Artificial Intelligence to Simulation-based Education: A Study of Comparative Educational and Operational Benefits. 2024. PubMed PMID: rayyan-206464992.
41. Domrös-Zoungrana D, Rajaeean N, Boie S, Fröling E, Lenz C. Medical Education: Considerations for a Successful Integration of Learning with and Learning about AI. Journal of Medical Education and Curricular Development. 2024;11:23821205241284719. PubMed PMID: rayyan-206464995.
42. Fatima SS, Sheikh NA, Osama A. Authentic assessment in medical education: exploring AI integration and student-as-partners collaboration. Postgraduate Medical Journal. 2024;100(1190):959-67. PubMed PMID: rayyan-206464997.
43. Galdames IS. From Anatomy to Algorithm: Scope of AI-Assisted Diagnostic Competencies in Health Sciences Education. International Journal of Medical and Surgical Sciences,(IJMSS). 2024;11(3):1-24. PubMed PMID: rayyan-206464999.
44. Huang H-S, Lee B-O. The Effectiveness of Applying Artificial Intelligence in Pediatric Clinical Education. 2024. PubMed PMID: rayyan-206465004.
45. Moldt J-A, Festl-Wietek T, Fuhl W, Zabel S, Claassen M, Wagner S, et al. Assessing Artificial Intelligence Awareness and Identifying Essential Competencies: Insights from Key Stakeholders in Integrating Ai into Medical Education. Available at SSRN 4713042. 2024. PubMed PMID: rayyan-206465015.
46. Roy AD, Das D, Mondal H. Efficacy of ChatGPT in solving attitude, ethics, and communication case scenario used for competency-based medical education in India: a case study. Journal of education and health promotion. 2024;13(1):22. PubMed PMID: rayyan-206465017.
47. Sridharan K, Sequeira RP. Artificial intelligence and medical education: application in classroom instruction and student assessment using a pharmacology & therapeutics case study. BMC Medical Education. 2024;24(1):431. PubMed PMID: rayyan-206465019.
48. Saleem N, Mufti T, Sohail SS, Madsen DØ. ChatGPT as an Innovative Heutagogical Tool in Medical Education. Cogent Education. 2024;11(1). doi: 10.1080/2331186X.2024.2332850. PubMed PMID: rayyan-206902202.

49. Gin BC, Cate Ot, O'Sullivan PS, Boscardin C. Assessing Supervisor versus Trainee Viewpoints of Entrustment through Cognitive and Affective Lenses: An Artificial Intelligence Investigation of Bias in Feedback. Advances in Health Sciences Education. 2024;29(5):1571-92. doi: 10.1007/s10459-024-10311-9. PubMed PMID: rayyan-206902211.
50. Soulage CO, Coppenolle FV, Guebre-Egziabher F. The Conversational AI "ChatGPT" Outperforms Medical Students on a Physiology University Examination. Advances in Physiology Education. 2024;48(4):677-84. doi: 10.1152/advan.00181.2023. PubMed PMID: rayyan-206902213.
51. Liu JQJ, Hui KTK, Al Zoubi F, Zhou ZZX, Samartzis D, Yu CCH, et al. The great detectives: humans versus AI detectors in catching large language model-generated medical writing. International Journal for Educational Integrity. 2024;20(1):8.
52. Ruf V, Dinc Y, Küchemann S, Berndt M, Steinert S, Kugelmann D, et al. Comparison and AI-Based Prediction of Graph Comprehension Skills Based on the Visual Strategies of First-Year Physics and Medicine Students. Physical Review Physics Education Research. 2024;20(2). doi: 10.1103/PhysRevPhysEducRes.20.020138. PubMed PMID: rayyan-206902233.
53. Chaiban T, Nahle Z, Assi G, Cherfane M. The Intent of ChatGPT Usage and Its Robustness in Medical Proficiency Exams: A Systematic Review. Discover Education. 2024;3. doi: 10.1007/s44217-024-00332-2. PubMed PMID: rayyan-206902234.
54. Kerth J-L, Bosse HM. The Future of Postgraduate Training? An Update on the Use of Entrustable Professional Activities in Pediatric Professional Training. Academic Pediatrics. 2024;24(7):1035-7.
55. Han Y, Choudhry HS, Simon ME, Katt BM. ChatGPT's Performance on the Hand Surgery Self-Assessment Exam: A Critical Analysis. Journal of Hand Surgery Global Online. 2024;6(2):200-5. doi: 10.1016/j.jhsg.2023.11.014. PubMed PMID: rayyan-207231744.
56. Landau M, Comeaux M, Mortell T, Boyle R, Imbrescia K, Chaffin AE. Characterizing the untapped potential of virtual reality in plastic and reconstructive surgical training: A systematic review on skill transferability. JPRAS Open. 2024;41:295-310. doi: 10.1016/j.jpra.2024.06.015. PubMed PMID: rayyan-207231770.

57. Guo FC, Chen YT, Hsu WC, Wang PC, Chen M, Chen JW. EMYWAY Workplace-Based Entrustable Professional Activities Assessments in Otolaryngology Residency Training: A Nationwide Experience. Otolaryngology–Head and Neck Surgery. 2025;172(4):1242-53.
58. Johnson D, Chopra S, Bilgic E. Exploring the Use of Natural Language Processing to Understand Emotions of Trainees and Faculty Regarding Entrustable Professional Activity Assessments. J Grad Med Educ. 2024;16(3):323-7. doi: 10.4300/JGME-D-23-00526.1. PubMed PMID: rayyan-206913226.
59. Wu Y, Zheng Y, Feng B, Yang Y, Kang K, Zhao A. Embracing ChatGPT for Medical Education: Exploring Its Impact on Doctors and Medical Students. JMIR Medical Education. 2024;10. doi: 10.2196/52483. PubMed PMID: rayyan-207231722.
60. Cherif H, Moussa C, Missaoui AM, Salouage I, Mokaddem S, Dhahri B. Appraisal of ChatGPT's Aptitude for Medical Education: Comparative Analysis With Third-Year Medical Students in a Pulmonology Examination. JMIR Medical Education. 2024;10. doi: 10.2196/52818. PubMed PMID: rayyan-207231723.
61. Luordo D, Torres Arrese M, Tristán Calvo C, Shani Shani KD, Rodríguez Cruz LM, García Sánchez FJ, et al. Application of Artificial Intelligence as an Aid for the Correction of the Objective Structured Clinical Examination (OSCE). Applied Sciences (2076-3417). 2025;15(3). PubMed PMID: rayyan-206465010.
62. Adarkwah MA, Badu SA, Osei EA, Adu-Gyamfi E, Odame J, Schneider K. ChatGPT in Healthcare Education: A Double-Edged Sword of Trends, Challenges, and Opportunities. Discover Education. 2025;4. doi: 10.1007/s44217-024-00393-3. PubMed PMID: rayyan-206902217.
63. Zhou Z, Rizwan A, Rogoza N, Chung AD, Kwan BYM. Differentiating between GPT-generated and human-written feedback for radiology residents. Current Problems in Diagnostic Radiology. 2025. doi: 10.1067/j.cpradiol.2025.02.002. PubMed PMID: rayyan-207231730.
64. Kang AJ, Rodrigues T, Patel RV, Keswani RN. Impact of Artificial Intelligence on Gastroenterology Trainee Education. Gastrointestinal Endoscopy Clinics of North America. 2025;35(2):457-67. doi: 10.1016/j.giec.2024.12.008. PubMed PMID: rayyan-207231716.

65. Ennab F, Farhan H, Zary N. Generative Artificial Intelligence and Its Role in the Development of Clinical Cases in Medical Education: A Scoping Review Protocol. 2025. PubMed PMID: rayyan-206464996.
66. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLoS digital health. 2023;2(2):e0000198.
67. Sallam M, Salim NA, Barakat M, Al-Mahzoum K, Al-Tammemi AaB, Malaeb D, et al. Assessing Health Students' Attitudes and Usage of ChatGPT in Jordan: Validation Study. JMIR Med Educ. 2023;9:e48254. doi: 10.2196/48254.
68. Matsieli M, Mutula S. COVID-19 and Digital Transformation in Higher Education Institutions: Towards Inclusive and Equitable Access to Quality Education. Education Sciences [Internet]. 2024; 14(8).
69. Wang S, Yang L, Li M, Zhang X, Tai X. Medical Education and Artificial Intelligence: Web of Science-Based Bibliometric Analysis (2013-2022). JMIR Med Educ. 2024;10:e51411. Epub 2024/10/10. doi: 10.2196/51411. PubMed PMID: 39388721; PubMed Central PMCID: PMCPMC11486481.
70. Goldenberg MG. Surgical Artificial Intelligence in Urology: Educational Applications. Urologic Clinics. 2024;51(1):105-15. doi: 10.1016/j.ucl.2023.06.003. PubMed PMID: rayyan-206465001.
71. Wang X, Liu XQ. Potential and limitations of ChatGPT and generative artificial intelligence in medical safety education. World J Clin Cases. 2023;11(32):7935-9. Epub 2023/12/11. doi: 10.12998/wjcc.v11.i32.7935. PubMed PMID: 38073698; PubMed Central PMCID: PMCPMC10698415.
72. Yang Y, Lin M, Zhao H, Peng Y, Huang F, Lu Z. A survey of recent methods for addressing AI fairness and bias in biomedicine. J Biomed Inform. 2024;154:104646. Epub 2024/04/28. doi: 10.1016/j.jbi.2024.104646. PubMed PMID: 38677633; PubMed Central PMCID: PMCPMC11129918.
73. Morley J, Floridi L, Kinsey L, Elhalal A. From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. Science and Engineering Ethics. 2020;26(4):2141-68. doi: 10.1007/s11948-019-00165-5.

Table 1. Inclusion and exclusion criteria for study selection

CATEGORY	INCLUSION CRITERIA	EXCLUSION CRITERIA
Publication type	Peer-reviewed journal articles	Conference abstracts, non-peer-reviewed sources, unpublished theses, editorials, commentaries, reports, preprints, or grey literature
Publication period	Published between 2010 and April 2025	Published before 2010 or after the search end-date; articles with unclear publication year
Language	English-language publications	Non-English articles, bilingual reports without English full-text
Full-text availability	Full-text articles available	Abstract-only articles, inaccessible full-text, or studies only available as protocol/registration
Focus and scope	Studies focusing on AI applications in medical education, especially within CBME frameworks	Studies with a primary focus on non-medical education, non-healthcare professions, or on technologies unrelated to AI
Educational context	Studies related to undergraduate, postgraduate, or continuing medical education using CBME approaches	Studies limited to high school, non-medical education, or medical training outside CBME/competency-based frameworks
Ai applications	Research involving machine learning, deep learning, intelligent tutoring systems, or generative AI (e.g., ChatGPT)	Studies focusing exclusively on non-AI digital technologies (e.g., traditional e-learning, telemedicine without AI, simulation without AI)
Competency focus	Studies presenting data on AI use for assessing cognitive, clinical, or non-cognitive competencies in medical students	Studies only describing AI for administrative tasks, resource management, or technical infrastructure, without relation to competency assessment

Table 2. Temporal trends and characteristics of AI research in medical education

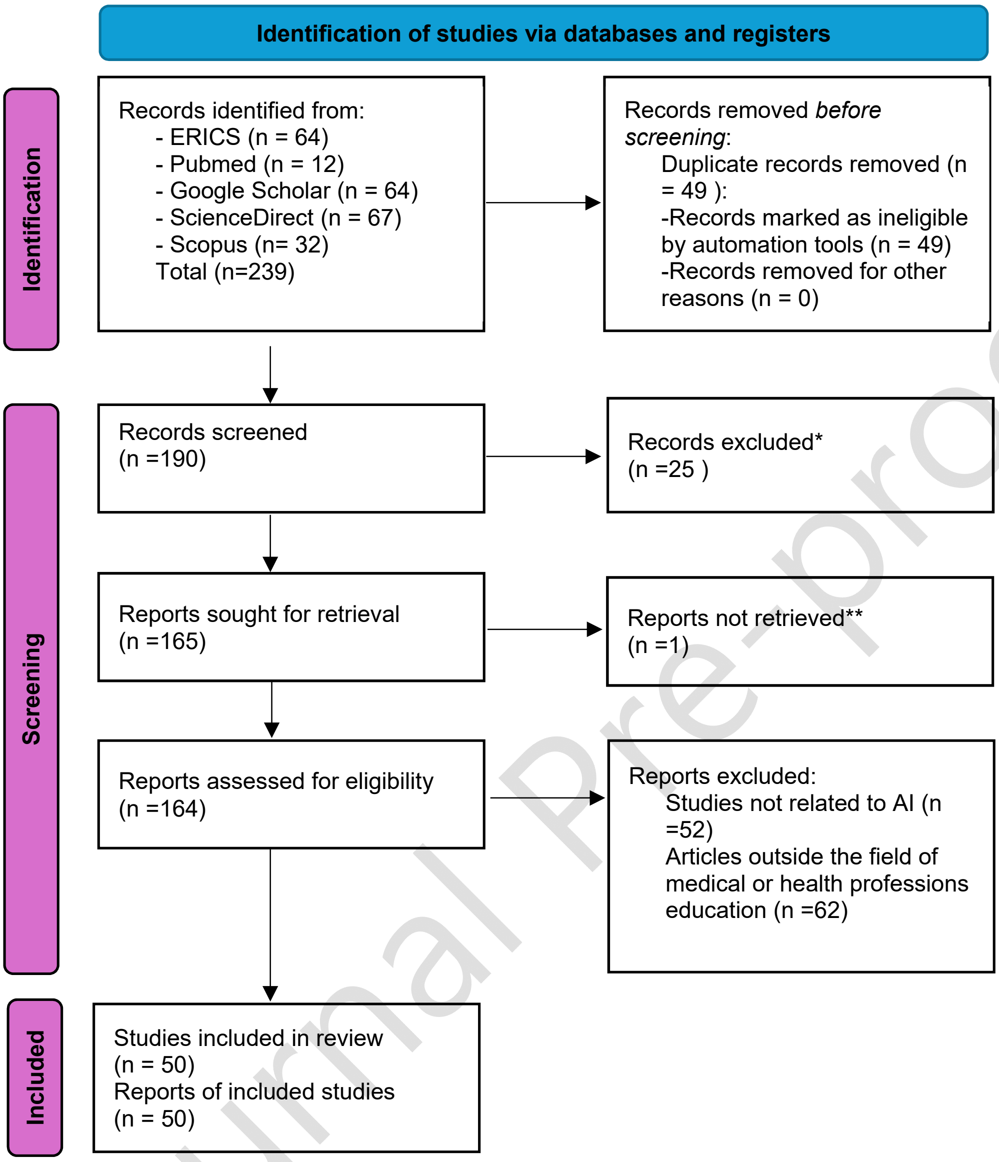
Year	Count	Countries	AI Types	Study Designs	Application Areas
2010 [18]	1	USA	Machine learning / deep learning	Quantitative study	Medical school basic science curriculum assessment
2019 [19, 20]	2	UK; USA	Machine learning / deep learning; ontology-based rule-driven system	Conceptual / perspective paper; tool development and evaluation study	Higher education - credentialing; curriculum innovation; medical education
2020[21]	1	Canada	Machine learning / deep learning	Review	Surgical video assessment
2021[22-25]	4	Canada; UK; USA; China	Intelligent tutoring system (ITS); learning analytics (LA); educational data mining (EDM); process mining; text analytics; machine learning / deep learning; optimization algorithm (metaheuristics; ML-based)	Review; exploratory case study; perspective / commentary; simulation-based modeling (Monte Carlo) + secondary data analysis	Medical Education
2022[26, 27]	2	Canada; Ukraine	AI in general; clinical decision support systems; murmur analysis; ECG interpretation; chatbot-based automation	Review	Medical education; pharmaceutical chemistry education
2023[28-37]	10	USA; Malaysia; Pakistan; Canada; India;	Natural language processing (NLP); learning analytics; virtual simulations; individualized feedback systems; artificial	Perspective/conceptual paper; qualitative study; cross-sectional study; review; experimental study; descriptive study of program	Medical education; surgical education; medical biochemistry education; nursing education;

		Saudi Arabia; India; Belgium; Australia; Thailand; China; Cambodia; Vietnam; Austria; Poland; Singapore; UK	intelligence; generative AI / LLM; machine learning / deep learning; AI-based automated essay scoring (AES); conversational agent; AI-based system	development and evaluation through a student survey; system architecture development and heuristic evaluation; viewpoint/commentary	medicine and public health in mics
2024 [14, 38-60]	26	USA; France; Germany; Pakistan; Chile; Taiwan; Australia & Canada; India; Bahrain; Norway; Hong Kong SAR; China; Lebanon; Canada;	Generative AI / LLM; AI; natural language processing (NLP); machine learning / deep learning; natural language processing; not specified; includes dashboard analytics and feedback tools; lexicon-based sentiment analysis	Perspective/conceptual paper; program description / educational innovation; experimental study; perspective / theoretical framework; conceptual / commentary; review; qualitative study; cross-sectional study; conceptual study with simulated AI-human interaction analysis; retrospective document analysis with AI-assisted thematic and sentiment analysis; quantitative study; performance evaluation study; sentiment analysis	Nursing education; digital health education; medical education; health sciences education – diagnostic training; cardiovascular medicine education; feedback analysis in clinical education; detection of AI-generated medical writing; graph comprehension and education research; postgraduate pediatric education; surgical training; plastic and reconstructive surgery; otolaryngology residency training; education (graph

		China; Tunisia		using NLP on survey text responses	comprehension skills assessment)
2025 [61-64]	4	United Arab Emirates; Spain; Germany; UK; USA; Ghana; Canada.	Generative AI / LLM; machine learning / deep learning	Review; experimental study	Medical education; clinical competency assessment (OSCE); healthcare education (medical; nursing; public health; dental; audiology; pediatric education); radiology education; endoscopy education

Table 3. Major thematic topics in AI applications within medical education

TOPIC	Number of Studies	TOP KEYWORD	Top Applications
Topic 1	2	ChatGPT, learning, language, large, machine, model, deep, LLM, processing, natural	Medical Education in Ophthalmology, Training Medical Students; Competency-Based Medical Education
Topic 2	10	learning, machine, large, generative, model, ChatGPT, vision, and, systems, SVM	clinical decision support, AI-enhanced curriculum design, and assessment frameworks
Topic 3	9	machine, language, ChatGPT, LLM, learning, analytics, systems, model, support, general	personalized learning systems, feedback analysis, interactive learning environments
Topic 4	7	learning, generative, language, LLM, large, model, general, deep, artificial, systems	Simulation-based education, intelligent tutoring systems, virtual healthcare environments
Topic 5	22	language, model, NLP, LLM, large, ChatGPT, GPT, systems, computer, processing	Automated essay scoring, diagnostic decision-making, patient feedback systems



*
– Abstract not related to AI (n = 13)
– Abstract not related to CBME (n = 9)
– Abstract not related to both AI and CBME (n = 3)

**
– Records not retrieved due to full-text inaccessibility (n = 1)

Figure 1. The PRISMA 2020 flow diagram of the scoping review

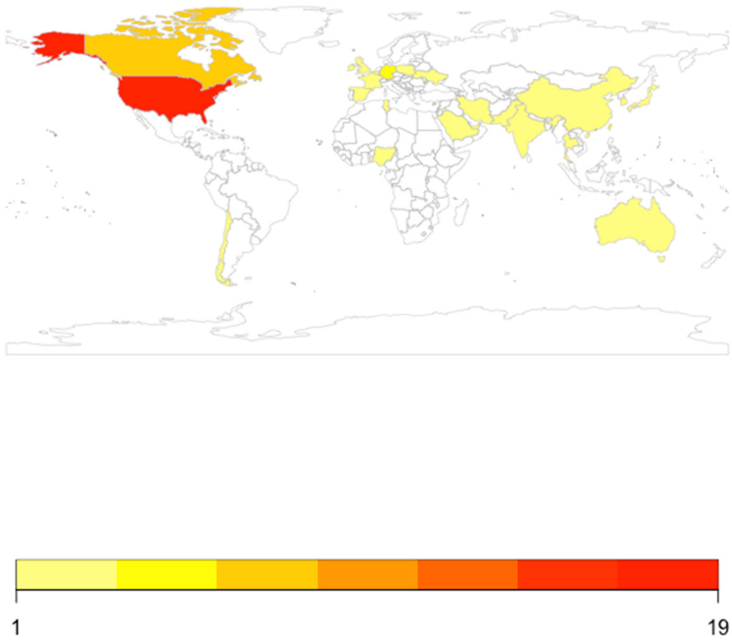


Figure 2. Global research output intensity in AI applications for medical education, visualized by country (latest updated on April 2025).

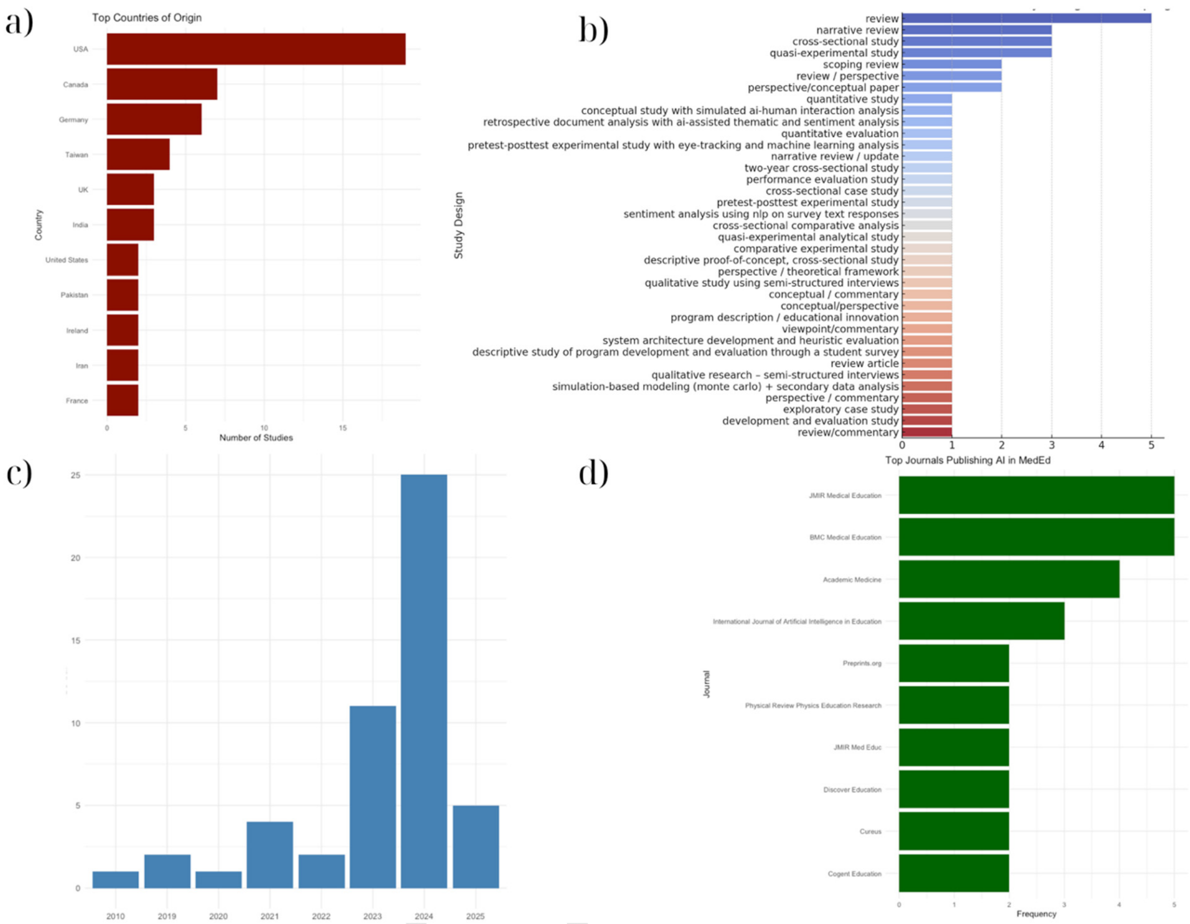
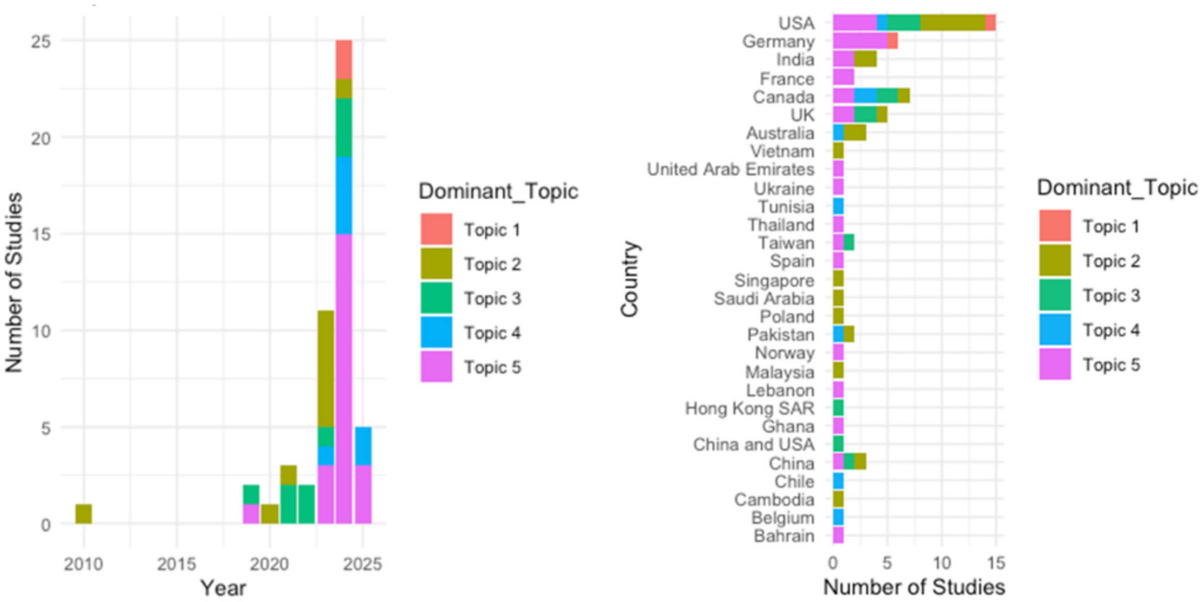


Figure 3. Overview of AI research in medical education: geographic origins, study designs, publication trends, and key journals

(a) Top countries of origin for AI in medical education research
(b) Distribution of study designs in included articles
(c) Number of publications per year
(d) Leading journals publishing AI research in medical education



Leftt Right

Figure 4. **Left**: Topic trends over time (latest updated on April 2025); **Right**: Distribution of topics by country

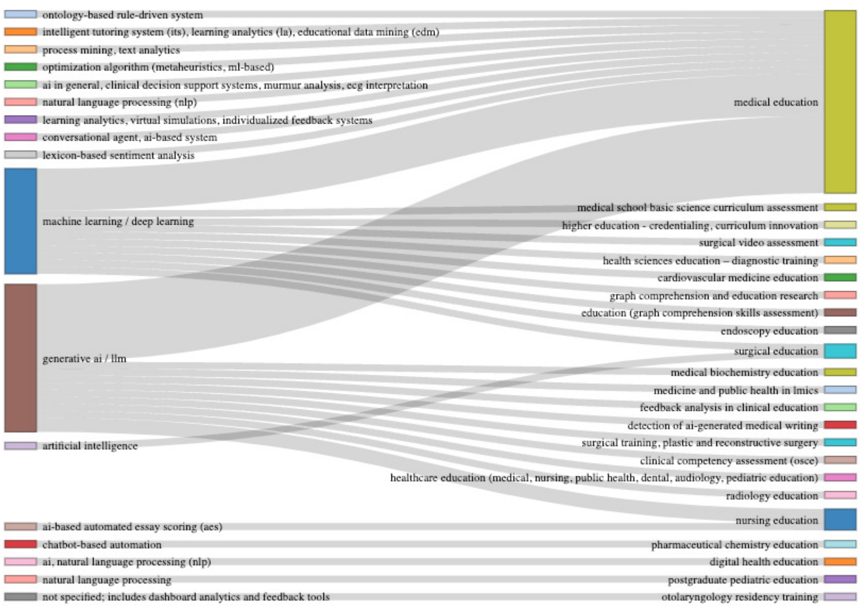


Figure 5. Sankey diagram showing flow from AI types to application areas

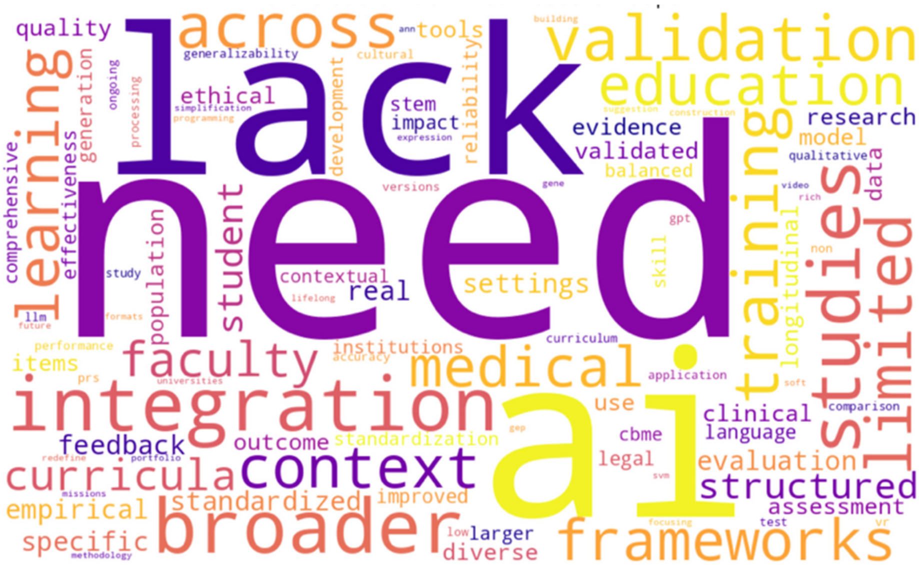


Figure 6. World cloud of gaps identified in the literature on AI in CBME