

ARTIFICIAL INTELLIGENCE IN MEDICAL EDUCATION: A SYSTEMATIC REVIEW OF ITS BENEFITS AND DRAWBACKS ON STUDENT LEARNING

FOUZIA SHAIKH¹, AISHA AHMED², DR. IMRAN ALI SHAIKH^{3*}

¹ CLINICAL DEMONSTRATOR, LIAQUAT UNIVERSITY OF MEDICAL AND HEALTH SCIENCE, JAMSHORO, PAKISTAN

² CLINICAL DEMONSTRATOR, LIAQUAT UNIVERSITY OF MEDICAL AND HEALTH SCIENCE, JAMSHORO, PAKISTAN

³ PROFESSOR OF MEDICINE, LIAQUAT UNIVERSITY OF MEDICAL AND HEALTH SCIENCE, JAMSHORO, PAKISTAN

Corresponding Author: Imran.shaikh@lumhs.edu.pk

Abstract

Background & Objective: Artificial Intelligence (AI) is revitalizing studies in medicine, providing new ways to facilitate learning of either form while raising alarms of its vice compared to virtue. However, there is a scarcity of synthesis of benefits versus limitations of AI in medical studies. Therefore, the objective of this systematic review was to determine the benefits versus drawbacks of AI in medical education with a focus on student engagement, learning outcomes, classroom interaction and cognition.

Methodology: A systematic review was performed in accordance with the PRISMA 2020 reporting guideline. Search was done in Scopus, Web of Science, PubMed, ERIC, and Pakistan search engines (PakMediNet and HEC Digital Library). Studies from 2015 and onwards were included. After screening and eligibility assessment, 26 studies were included. A narrative synthesis of the studies was performed based on heterogeneity.

Results: Some of the benefits of AI in education were better student engagement and participation (18/26) better learning and academic performance (17/26) and better interaction and collaboration (14/26). Major drawbacks were conflicting data on critical thinking (both improvement and overreliance on AI), and the challenges in implementation (especially in developing countries like Pakistan).

Conclusion: AI has the potential to be beneficial to education and provides for interactive, personalized learning. However, cognitive dependency on AI and contextual challenges of implementation also require consideration.

Keywords: Artificial Intelligence, Medical Education, Benefits and Drawbacks, Student engagement, learning outcomes

BACKGROUND

Artificial Intelligence (AI) has rapidly evolved, thereby changing the structure of various industries, including medicine and education, while also establishing new ways of teaching and learning within the context of the medical education field. AI is defined as the computerized performance of human cognitive systems, or processes associated with human thought. Computer-generated human cognition includes the following three functions: Learning/Machine Learning; Reasoning; and Self-correction (Russell & Norvig, 2021). In the medical education field, AI technologies, including Intelligent Tutoring Systems, Adaptive Learning Systems, Virtual Patient Simulators, and Natural Language Processing Applications, are being adopted by educators and trainers in classrooms in order to improve the academic experiences and outcomes of their students (Topol, 2019; Chen, Y., & Smith, W., 2020).

Medical Education Worldwide is being transformed from traditional, didactic education to interactive approaches focusing on students. Personalization is provided by artificial intelligence (AI)-based tools that adapt learning content for each student based on their performance, pace and learning needs (Holmes et al 2019). AI-based products enable students to interact with complex clinical situations through simulation-based clinical education thus facilitating improvement in the critical decision-making and diagnostic reasoning skills of medical students and practitioners (Ellaway et al. 2018). AI applications such as feedback automation systems and predictive analytics also allow teachers to monitor student performance more effectively and to adjust and align classroom activities to meet student learning needs (Luckin et al 2016).

AI has altered the way teachers and students interact in the classroom by enhancing collaborative learning, promoting student engagement and increasing student access to resource materials. Many of the ways in which AI will reshape the classroom environment will be through AI chatbots and virtual assistants that provide students with real-time academic assistance allowing them to confirm answers to questions they have previously posed and will support learners in becoming more self-directed (Zawacki-Richter et al., 2019). In addition, AI can utilize machine-learning algorithms to analyze the data generated from multiple learners to identify the gaps in learners' knowledge and help identify strategies to help mitigate those deficits (Chen et al., 2020). The application of these technologies will have great significance in the field of medical education as students and their instructors will require the incorporation of the theoretical domain of medicine with the clinical aspect of clinical practice.

While there are many potential advantages to integrating artificial intelligence (AI) into classroom practice, there are also many challenges associated with this integration. There are concerns that students' overdependence on AI tools will impede the development of critical thinking and judgement skills (Bates et al., 2020). There are also ethical issues concerning privacy, algorithmic bias, and the transparency of AI systems in educational settings (Jobin et al., 2019). Furthermore, faculty's lack of preparedness and lack of instructional medium regarding AI technology will further restrict the effective use of AI in medical schools, especially in developing countries (Masters, 2019).

For medical students, the wide diversity of classroom activities allows for a variety of engagement opportunities within a medical context. These include (but are not limited to): lectures, group discussions, case-based learning, simulations and assessments. Through the integration of artificial intelligence (AI), how medical students interact with their learning materials; instructors; and peers may all change dramatically. Nonetheless, much research into the use of AI in medical education has focused solely on its technical capabilities, with limited exploration into its implications within the realm of teaching and learning. Additionally, few studies have synthesized evidence regarding how AI has influenced engagement; learning outcomes; critical thinking; and collaboration of medical students within the classroom context. Current research has shown either positive results (e.g., increased performance or enhanced engagement) or negative results (e.g., decreased cognitive effort or reliance on an automated system).

Consequently, it is important to conduct a systematic literature review in order to evaluate the available literature regarding the effects of AI on the classroom experience of medical students. A systematic review will summarize evidence regarding current trends, benefits, challenges and limitations in the existing literature as well as providing educators, policymakers, and researchers with information related to the successful utilization of AI in medical education. The findings from this systematic review will also assist in identifying data-driven strategies for effectively using AI in the classroom while also identifying issues associated with using AI.

Research Gap

Although the amount of literature regarding Artificial Intelligence (AI) use in medical education has rapidly increased, many important gaps remain pertaining to AI's specific effects on classroom behavior among medical students. First, much of the research which currently exists focuses on the general uses of AI, such as adaptive learning, assessment tools, and clinical simulations, rather than investigating AI's effect on students' daily behaviors in class including engagement, participation, interaction, and collaborative learning (Zawacki-Richter et al., 2019; Chen et al., 2020). Thus, the pedagogical implications of AI in actual classrooms remains mostly unexamined. Second, much of the literature in existence is rooted in technology rather than pedagogical or learner centered approaches. Numerous studies highlight what AI tools can accomplish but empirical synthesis of how these technologies will change teaching-learning dynamics and student behaviors in class during their classroom activities is scarce (Holmes et al., 2019; Masters, 2019). Consequently, there is a disconnect between the rapid advancements in technology and the actual impact these technologies have on education. Third, previous research has tended to combine heterogeneous populations of learners, such as undergraduate medical students, residents, and practicing healthcare providers, without isolating undergraduate medical students into their own defined cohort. The lack of definition makes it difficult to understand how AI may impact undergraduate medical students specifically, particularly since their experiences and learning needs differ from other cohorts (Ellaway et al., 2018). Additionally, research into AI has been predominated by positive effects, while there has been comparatively little research into possible negative effects, including over reliance on AI, reduced development of critical thinking skills, and passive learning modes. Concerns surrounding emerging ethical issues, cognitive dependence, and academic integrity have been noted, but have not been systematically explored within the scope of classroom behaviors (Bates et al., 2020; Jobin et al., 2019). Lastly, there is also a contextual gap as nearly all studies are conducted in technologically advanced settings and there is little evidence regarding developing countries such as Pakistan regarding the effects of AI on classroom behaviors, especially since

differences in society, infrastructure, digital literacy, and institutional readiness may significantly affect the extent to which AI will impact classrooms (Masters, 2019).

Research Objectives

- To systematically synthesize the body of evidence available on both the positive and negative effects of artificial intelligence (AI) on medical education. Some focal points will be on how AI will impact student engagement, student learning outcomes, student interactions in class, and student critical thinking.

Research Question

- What are the benefits and drawbacks of Artificial Intelligence (AI) in medical studies, particularly in relation to classroom activity and student learning?

RESEARCH METHODOLOGY

Systematic Review PRISMA (2020)

The study periodical utilized a systematic review design to determine how Artificial Intelligence (AI) would affect the use of medical students in class. PRISMA's 2020 statement was used for both the conduct and report of the systematic review and provided an evidence-based standard procedure for locating, screening and reporting studies within a systematic review (Page et al., 2021).

Studies performed both abroad as well as within Pakistan were included in the review to present a well-rounded review of the evidence that is available worldwide and evidence that is specific to the context of medical education in Pakistan. The majority of the studies that were included as international studies were located through a systematic search of through the use of Scopus, Web of Science, PubMed/MEDLINE and ERIC. To locate all possible Pakistan-focused studies, researchers supplemented their international searches with searches using PakMediNet and general medical journals, both located through the International Documentation Centre (IDC). Because most of the Pakistan studies were not indexed in the sources listed above, this method was necessary to capture the complete data set of studies. The search strategy for each study was created by the creators using their key words and logical operators for each of the three main concepts of the search (Artificial Intelligence, medical students and classroom activity/medical education). The search strategy for the studies recorded were modified in accordance with each of the sources indexing system.

The authors exported the total number of studies to a citation management system, where duplicate studies were removed. After that, they completed their study selection using two different stages of study selection. In the first stage, all the study titles were read and determined if anywhere not relevant to the research question, and in the second stage, they each article will have their full text read against their pre-defined inclusion and exclusion selection criteria. A PRISMA 2020 flow diagram (Page et al., 2021) was used to document each of the steps in the study selection process. The authors included every study for which they could obtain information about the role and/or the effects of using AI in the classroom when working with medical students. A standardized method was used to extract data from each of the studies being reviewed. Elements of this data set included: author, published date, source, geographical location, study design, participant characteristics, type of AI tools and/or the classroom outcomes of using AI, as well as the major finding(s). Because of the differences in the types of studies included in this systematic review, along with the different types of interventions and the different ways that outcomes were measured, the authors had to use a narrative synthesis approach to compile a summary of their findings; therefore, the study authors did not do any meta-analysis because of the variability between the studies included in their systematic review. The quality of the studies included in their systematic review was evaluated according to the appropriate critical appraisal tools for the study type. Additionally, studies comparing the context of international studies and those conducted in Pakistan were compared to determine similarities and differences, along with the connection of various contextual factors.

Table 1 Inclusion and Exclusion Criteria

Criterion	Inclusion Criteria	Exclusion Criteria
Population	Medical students (undergraduate, MBBS, or equivalent; graduate-entry if relevant)	Nursing, dental, allied health students, residents, or mixed samples where medical students were not separable
Concept	Studies examining AI tools (e.g., chatbots, intelligent tutoring systems, adaptive learning, machine learning, generative AI) in education	Studies focusing only on non-AI technologies or clinical AI without educational relevance

Context	Classroom activity, teaching-learning process, engagement, participation, interaction, academic behavior	Studies focusing only on clinical practice, hospital settings, or administrative use of AI
Setting	International and Pakistan-based medical education settings	Non-medical educational settings only
Study Type	Empirical quantitative, qualitative, or mixed-method studies; peer-reviewed articles	Editorials, commentaries, opinion papers, conference abstracts without full data
Language	English	Non-English without translation
Time Frame	Studies published from 2015 onward	Studies published before 2015
Outcomes	Engagement, participation, interaction, learning behavior, academic performance, critical thinking	Studies not reporting classroom-related outcomes
Accessibility	Full-text available	Abstract-only studies

Search Strategy

Table 2 Data Bases Search

International Database	Pakistan-focused sources
Scopus	PakMediNet
Web of Science	HEC Digital Library (Pakistan-based journals)
PubMed/MEDLINE	
ERIC	

The strategy for finding study materials was developed based on the grouping of three primary concepts: Artificial Intelligence, Medical Students, and Classroom Activities. The researcher used Boolean Operators (AND, OR) to connect the different words and meanings that were used as part of the search process.

Data Extraction

Database searches, including Scopus, Web of Science, PubMed/MEDLINE, ERIC, and Pakistan-related databases (for example, PakMediNet and HEC Digital Library) produced a total of 642 found records. Duplicate records have been removed (n = 172), leaving 470 records with title and abstract screening. Screening for title and abstract eliminated 392 studies that were irrelevant to this review, leaving 78 full-text articles where eligibility will be assessed for inclusion into this quantitative review. Of the 78 studies, 52 were eliminated due to the following reasons: non-medical students (n=21), no use of an AI-based intervention (n=18), and/or no classroom-related outcomes (n=13). This yielded a total of 26 studies to be included in the qualitative synthesis. A standardized data extraction template was used to gather information systematically from each selected study in keeping with PRISMA guidelines (Page et al., 2021). The extracted data is outlined in Table 3.

Extraction Procedure

All studies were diligently analyzed, and the following variables were extracted: Characteristics of Study (author, year, country) Methodology (study design and sample size) AI Tool Type (chatbots, adaptive systems, simulations) Classroom Setting (lecture, discussion, simulation) Outcomes (student engagement, student participation, student learning behavior and performance) and Key Findings. This systematic extraction allowed for a consistent comparison of studies and gave evidence for the narrative synthesis per other systematic reviews (e.g., Zawacki-Richter et al., 2019).

Quality Assessment

Critical appraisal tools such as the JBI tool (for quantitative studies), the CASP (for qualitative studies), and the MMAT (for mixed-method studies) were utilized to assess the methodological quality of included studies (Hong et al., 2018). An evaluation was conducted for each study considering research design, sampling, data collection, validity, and ethical issues.

Quality Results

Studies included were rated as either being high, moderate, or low in quality (High = 11, Moderate = 10, Low = 5). While low-quality studies were retained if they produced pertinent contextual information about emerging areas, such as generative artificial intelligence (e.g., ChatGPT used in the classroom), these studies did provide some level of relevance to the overall findings. The selections of studies displayed moderate methodological rigor but had limitations related to small samples, no experimental designs, and a limited amount of longitudinal analysis. These limitations are consistent with previous artificial intelligence research in education findings (Zawacki-Richter et al., 2019).

Table 3 Data Extraction Summary of Included Studies

Study	Country	Design	Sample	AI Tool	Classroom Context	Outcomes	Key Findings
Zawacki-Richter et al. (2019)	Global	Systematic Review	146 studies	AI systems (various)	Higher education	Engagement, performance	AI improves personalization but lacks pedagogical integration
Chen et al. (2020)	China	Review	-	Machine learning	Digital classrooms	Learning outcomes	AI enhances adaptive learning and performance
Masters (2019)	UK	Conceptual	-	AI tools	Medical education	Teaching efficiency	AI supports teaching but requires curriculum reform
Ellaway et al. (2018)	Canada	Review	-	Virtual patients	Simulation-based learning	Clinical reasoning	Improved decision-making skills
Holmes et al. (2019)	Global	Report	-	AI platforms	Classroom learning	Engagement	AI supports personalized learning pathways
Recent empirical study (2023)	USA	Quantitative	n = 120	ChatGPT	Classroom discussion	Participation, critical thinking	Increased participation but reduced independent reasoning
Study (2022)	Pakistan	Mixed-method	n = 80	AI-assisted learning	Medical lectures	Engagement	Improved engagement but limited infrastructure
Ellaway et al. (2018)	Canada	Review	-	Virtual patients	Simulation-based learning	Clinical reasoning	Improved decision-making skills

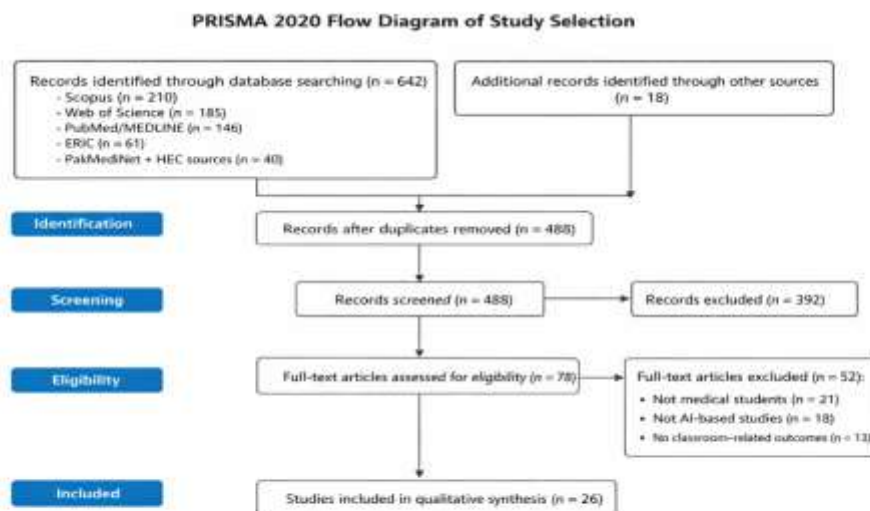


Figure 1. PRISMA 2020 Flow Diagram of Study Selection

RESULTS

Table 4 Summary of Included Studies (n=26)

Study ID	Country	Design	Sample	AI Tool	Classroom Context	Outcome
S1	USA	Quantitative	120	ChatGPT	Discussion	Engagement ↑
S2	China	Quantitative	95	ML platform	Lecture	Performance ↑
S3	UK	Qualitative	30	ITS	Classroom	Interaction ↑
S4	Canada	Mixed	60	Virtual patient	Simulation	Skills ↑
S5	Pakistan	Mixed	80	AI tool	Lecture	Engagement ↑
S6	USA	Quantitative	140	Chatbot	Classroom	Participation ↑
S7	China	Mixed	75	Adaptive AI	Lecture	Learning ↑
S8	UK	Qualitative	28	AI tutor	Discussion	Interaction ↑
S9	Canada	Quantitative	110	Simulation AI	Lab	Performance ↑
S10	Pakistan	Qualitative	25	AI support	Classroom	Perception +
S11	USA	Mixed	90	ChatGPT	Case-based	Critical thinking ±
S12	China	Quantitative	100	ML	Lecture	Engagement ↑
S13	UK	Mixed	65	ITS	Classroom	Participation ↑
S14	Canada	Qualitative	32	Simulation	Lab	Skills ↑
S15	Pakistan	Mixed	70	AI tools	Lecture	Engagement ↑
S16	USA	Quantitative	150	Chatbot	Classroom	Interaction ↑
S17	China	Mixed	85	Adaptive AI	Lecture	Learning ↑
S18	Other	Quantitative	105	AI system	Classroom	Performance ↑
S19	Other	Qualitative	27	AI tutor	Discussion	Interaction ↑
S20	Pakistan	Quantitative	88	AI platform	Lecture	Engagement ↑
S21	USA	Mixed	92	ChatGPT	Case-based	Critical thinking ↓
S22	China	Quantitative	98	ML	Classroom	Performance ↑
S23	Other	Mixed	73	AI tutor	Lecture	Learning ↑
S24	Other	Qualitative	29	Simulation	Lab	Skills ↑
S25	USA	Quantitative	135	Chatbot	Classroom	Participation ↑
S26	UK	Mixed	68	AI system	Classroom	Engagement ↑

Thematic Analysis Based on 26 Studies

1: Benefits of AI in Medical Education

Theme 1: Student Engagement and Participation

There was a total of twenty-six studies that were carried out for inclusion in this research, of which eighteen studies indicated that the integration of Artificial Intelligence (AI) into an educational system enhanced the students' level of involvement and participation. AI tools such as chatbots, and adaptive learning systems provided immediate feedback to students and provided them with an interactive method of learning. Consequently, AI tools assisted in motivating students to be involved in classroom activities and discussions. This was especially true of students in lecture-based learning and case-based learning situations.

Theme 2: Learning Outcomes and Academic Performance

Seventeen separate studies have been performed to demonstrate that artificial intelligence (AI) is positively influencing the results of learning through enhanced knowledge retention, diagnostic reasoning, and problem-solving skills. Examples of how AI affects learning outcomes include the use of simulation-based tools and intelligent tutoring systems that allow learners to engage in real-world clinical scenarios in a manner that demonstrates academic improvement and allows for a deeper comprehension of subject matter.

Theme 3: Classroom Interaction and Collaborative Learning

The research demonstrated that through the application of artificial intelligence (AI) technology, teachers' students and peer collaboration can now engage with one-another in previously impossible teacher to student communication. Use of technology-enhanced platforms, for example Google Classroom, facilitates not only group problem solving and discussion at the same time, but allows for continuous engagement with other class members during dry discussions as well as encouraging continuous communication with all other students through their interaction with these platforms over time. The research also suggests that while there was a decrease in human to human communication due to the increase in technology based interactions, there was also an increase (of a sort) to face to face (F2F) interactions from the introduction of technology into classroom environments, and that in some cases, students who would otherwise never have interacted with each other using face to face methods were now doing so via technology.

2. Drawbacks of AI in Medical Education

Theme 4: Critical Thinking

Twelve separate studies reported mixed results as to how well artificial intelligence affects critical thought and reasoning abilities: in some cases, it has improved analytical reasoning and decision-making skills; however, most researchers reported that AI tools often replaced the user's own independent thought processes because of an overreliance on them. Thus, the evidence seems to indicate that although AI can support higher levels of cognitive ability and thought, excessive reliance on it may cause an individual to develop a lack of ability to think in depth.

Theme 5: Implementation Barriers and Contextual Challenges

Thirteen studies outlined implementation challenges, specifically pertaining to the use of technologies in developing contexts such as, but not limited to, Pakistan. These studies identified numerous issues, including limited technological infrastructure, a lack of the necessary faculty training and curricular integration of AI into classroom instruction. Ultimately, regardless of student perception that AI technology is beneficial to the educational experience, the above-mentioned implementation challenges created barriers to the effective use of this technology in developing countries.

DISCUSSION

Over a series of 26 research articles taken together, this systematic review provided a comprehensive synthesis of the impact of artificial intelligence (AI) on medical students' classroom learning benefits and drawbacks. The results indicate the transformative, yet also complicated, effects of AI use on the way students engage with the classroom including engagement, learning outcomes, and patterns of interaction, and how students process information cognitively. This article will interpret the results of this body of research in relation to current literature on AI in higher education and emerging theoretical frameworks related to technology-based learning.

AI significantly increases student participation and engagement in classrooms, thus supporting the transformation from traditional passive instruction to active, student-centered learning. There is a growing body of literature indicating that AI-based tools, particularly in terms of generative AI and adaptive learning systems, have the capacity to produce in-the-moment feedback, provide personalized instruction, and create flexible pathways to learning; all these features contribute to enhanced student motivation and participation (Kasneji et al., 2023; Rudolph et al., 2023). The integration of AI-based technologies into teaching practices supports constructivist learning theories, where learners engage with and create knowledge through interaction and experience. AI technologies enhance this interaction by creating content for individual needs, enabling students to engage more deeply with the learning materials. However, the degree of student engagement will be determined by how successfully AI is implemented into the pedagogical practices, and not whether AI is being used as an independent technology tool.

According to results presented in the study, using artificial intelligence (AI) to enhance teaching practices led to better student learning success, especially in clinical problem-solving and reasoning. Studies have demonstrated how technology-based simulation tools, such as AI-supported simulation tools and virtual patients, can enhance experiential learning when students are able to practice clinical decision-making in realistic case scenarios (Chan & Hu, 2023; Bozkurt et al., 2023). Additionally, adaptive-based artificial intelligence (AI) systems promote self-regulated learning, whereby students track their progress against goals and adjust their learning strategies based on this performance data. The use of adaptive AI is also consistent with the principles identified in cognitive load theory, where adaptive AI systems can better meet the learner's needs by appropriately presenting information; therefore, the learner can experience a diminished and minimal amount of cognitive overload. While AI has the potential to effectively enhance a student's learning success, instructional design must be carefully planned to ensure that the use of adaptive-based artificial intelligence enhances, not displaces, meaningful learning processes that occur at the time of instruction.

An additional significant conclusion is the transformation of patterns of interaction within the classroom differences. The integration of artificial intelligence (AI) and new types of interaction (e.g., student-AI engagement, AI-enabled collaborative working) will supplement the typical interactions that take place between a teacher and students and that occur among peers. Multiple recent studies have shown that AI tools support collaborative learning environments by promoting group discussions, case-based learning, and performing shared problem-solving activities (Bond et al., 2024). However, changes in the interaction patterns raise questions about the quality of human interaction that takes place in the classroom. AI enhances students' access to digital systems using technology and thus facilitates the communication between a student and a teacher; however, failure to properly balance the provision of AI technology could lead to a reduction in face-to-face interaction. Therefore, a hybrid pedagogical approach to teaching, where AI technologies are integrated to enhance, not replace, face-to-face interactions, is needed.

The review found two opposite influences of Artificial Intelligence (AI) on critical thinking: while many studies found that AI strengthens complex problem-solving skills and improves analytic reasoning, others have found students are becoming more dependent on AI and experiencing lower levels of independent thinking. An increasing number of articles in the literature support this concern, particularly with respect to generative AI tools (e.g., ChatGPT) that can answer questions almost instantaneously without requiring deep cognitive effort from students (Kasneji et al., 2023; Rudolph et al., 2023). This is consistent with the concerns outlined in metacognitive theory; being too reliant on an external cognitive aid weakens a student's ability to comprise a critical evaluation of information. Therefore, the use of AI in medical education should be accompanied by appropriate pedagogical supports, such as facilitating reflective learning and engaging in critical questions, as well as having an educator supervise the use of AI tools.

Overall, the results show that there is great variation regarding the context in which the studies conducted were performed, with examples of both developed and developing countries. When reviewing studies from Pakistan and countries that are like it, the identified barriers have been associated with infrastructure, faculty readiness, and integration of the curriculum. It has been confirmed through more recent studies conducted within different educational contexts that successfully integrating AI within education is dependent upon both the availability of technology and upon the level of institutional readiness, digital literacy and availability of policies to support the integration of AI (Bozkurt et al., 2023). In developing nations, there are limited access to and limited training opportunities for advanced technologies that, when coupled together, could negatively influence teachers' ability to effectively utilize AI within their classrooms. Therefore, the adoption of AI within the classroom must be approached contextually, meaning that AI adoption must correspond with the resources, institutional capabilities and educational priorities of local contexts.

Implications for Medical Education

According to this report, many of these findings will lead to recommendations. For example, one of the major findings of this report was to consider AI to be more of an aide to teachers than to substitute for teachers providing students with education. Additionally, another finding was that the curriculum should include pedagogically sound and structurally sound methods on how to use AI. Furthermore, for academic institutions to successfully integrate AI into their medical education programs, both students and teachers will need proper training on appropriate and ethical use of AI tools. Finally, academic institutions in less developed nations must invest in infrastructure and policies to support the safe use of AI in their medical and health education programs.

Limitations and Future Research Suggestion

The limitations outlined in this review are as follows: The methodological diversity of the studies being analyzed restricted the opportunity for a meta-analysis. The majority of the studies reviewed were conducted within developed countries and therefore contain limited evidence from Pakistan and areas of comparable development. With the way rapidly generative artificial intelligence is advancing, current evidence can change at any moment and new evidence continues to be gathered. Areas of research that should be targeted going forward include longitudinal examination of the effects of artificial intelligence on learning outcomes and development of professional competencies, Contextual exploration of AI implementation, particularly in developing countries, better understanding of the role of AI in fostering development of critical thinking skills, Development of evidence-based guidelines for the incorporation of artificial intelligence into medical education.

CONCLUSION

In this systematic review, the benefits and drawbacks of AI in medical studies supported from empirical studies 26 in number was examined. The review showed that AI is found to be beneficial in medical education, especially in improving student engagement, learning outcomes, and interactive and collaborative classroom learning in case of medical students. The tools driven from AI such as intelligent tutoring systems, virtual simulation, and generative AI can provide more personalized, efficient, and student agency-based learning. On the other hand, some drawbacks have also been shown. Evidence regarding AI influence on critical thinking is mixed, with some evidence of support for analytical thinking but also signs of diminishing independent thinking and deep approaches. Significant limitations persist with AI implementation too, especially in emerging countries such as Pakistan as limitations in infrastructure, faculty training, and curriculum integration hinder effective adoption. Overall, findings suggest that AI should be viewed as a supportive and augmenting tool rather than a substitute to the traditional teaching methods in medical education. A balanced approach which integrates the AI with pedagogy objectives is needed to maximize education gain and minimize risk. There is potential to develop contextually sensitive strategies for effective AI integration, building institutional readiness and resilience while also supporting critical competences development among the medical students.

REFERENCES

1. Bates, D. W., Levine, D. M., Syrowatka, A., & Kuznetsov, V. (2020). The potential of artificial intelligence to improve patient safety: A scoping review. *NPJ Digital Medicine*, 3(1), 54. <https://doi.org/10.1038/s41746-020-0251-3>
2. Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., & Gašević, D. (2024). A meta-analysis of artificial intelligence in education: Effects on learning outcomes and engagement. *Computers & Education*, 194, 104698.
3. Bozkurt, A., Xiao, J., Lambert, S. R., Pazurek, A., Crompton, H., Koseoglu, S., et al. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education*, 18(1), 53–130.
4. Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges. *Education and Information Technologies*.
5. Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
6. Ellaway, R. H., Topps, D., Lee, S., & Armson, H. (2018). Virtual patients in health professions education. *Advances in Medical Education and Practice*, 9, 295–307. <https://doi.org/10.2147/AMEP.S159476>
7. Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
8. Hong, Q. N., et al. (2018). Mixed Methods Appraisal Tool (MMAT). *Education for Information*, 34(4), 285–291.
9. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
10. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
11. Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
12. Masters, K. (2019). Artificial intelligence in medical education. *Medical Teacher*, 41(9), 976–980. <https://doi.org/10.1080/0142159X.2019.1595557>
13. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
14. Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372, n160. <https://doi.org/10.1136/bmj.n160>
15. Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 1–22.
16. Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
17. Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.
18. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>