Applications and challenges of Artificial Intelligence impact on the academic performance of medical students: a Scoping Review

Vanessa Uberuaga, Juan Ignacio Armijo, Vicente Aitor Tagle, Óscar Jerez

Submitted to: JMIR Preprints
on: December 05, 2024

Table of Contents

Applications and challenges of Artificial Intelligence impact on the academic performance of medical students: a Scoping Review

Vanessa Uberuaga[1*]; Juan Ignacio Armijo[1*]; Vicente Aitor Tagle[1*]; Óscar Jerez[1*]

[1]University of Chile Santiago CL
[*]these authors contributed equally

Corresponding Author:
Vanessa Uberuaga
University of Chile
Avenida Independencia 1027, Independencia
Santiago
CL

Abstract

Background: Artificial intelligence (AI) is an emerging field that has had a profound impact on medical education. Its growing implementation offers personalized learning environments, promising improvements in the academic performance of medical students.

Objective: This study aims to explore how AI tools influence the academic performance of medical students by identifying key applications, benefits and challenges and providing an integrative perspective on their role in medical education.

Methods: A scoping review was conducted based on empirical studies, systematic reviews and case studies published between 2014 and 2024. Studies were identified through PubMed, Scopus, and Web of Science, using search strategies developed with AI assistance. Eligible studies involved medical students or trainees and evaluated the impact of AI on academic performance, simulations, and clinical skills development.

Results: Initially 173 studies were identified, of which 108 met the inclusion criteria after screening. The final analysis included eight studies. Key findings showed that AI-enhanced teaching methods significantly improved academic performance compared to traditional approaches. In addition, AI-based simulations and personalized learning platforms contributed to improved practical skills and knowledge retention.

Conclusions: AI has demonstrated considerable potential to enhance medical education by improving academic performance and practical skills. However, challenges remain, including the need for formal AI education in medical curricula and careful evaluation of the effectiveness of AI tools. Further research is needed to explore the role of AI in specific areas of medical education and ensure its appropriate integration orinto educational systems.

(JMIR Preprints 05/12/2024:69724)
DOI: https://doi.org/10.2196/preprints.69724

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original paper

Óscar Jerez[1] ; Juan Armijo[2] ; Vicente Tagle[2] ; Vanessa Uberuaga[2]

[1]Department of Health Sciences Education, School of Medicine, University of Chile, Santiago, Chile.
[2]School of Medicine, Faculty of Medicine, Universidad de Chile, Santiago, Chile.

Corresponding Author:
Óscar Jerez
Department of Health Sciences Education
University of Chile
Independencia 1027, Independencia
Santiago, RM, 8380453
Phone: +56990799169
Email: ojerez@uchile.cl

Applications and challenges of Artificial Intelligence impact on the academic performance of medical students: a Scoping Review

Abstract

Background: Artificial intelligence (AI) is an emerging field that has had a profound impact on medical education. Its growing implementation offers personalized learning environments, promising improvements in the academic performance of medical students.
Objective: This study aims to explore how AI tools influence the academic performance of medical students by identifying key applications, benefits and challenges and providing an integrative perspective on their role in medical education.
Methods: A scoping review was conducted based on empirical studies, systematic reviews and case studies published between 2014 and 2024. Studies were identified through PubMed, Scopus, and Web of Science, using search strategies developed with AI assistance. Eligible studies involved medical students or trainees and evaluated the impact of AI on academic performance, simulations, and clinical skills development.
Results: Initially 173 studies were identified, of which 108 met the inclusion criteria after screening. The final analysis included eight studies. Key findings showed that AI-enhanced teaching methods significantly improved academic performance compared to traditional approaches. In addition, AI-based simulations and personalized learning platforms contributed to improved practical skills and knowledge retention.
Conclusions: AI has demonstrated considerable potential to enhance medical education by improving academic performance and practical skills. However, challenges remain, including the need for formal AI education in medical curricula and careful evaluation of the effectiveness of AI tools. Further research is needed to explore the role of AI in specific areas of medical education and ensure its appropriate integration orinto educational systems.
Keywords: Artificial intelligence; Academic performance; Medical students; Medical education; Clinical simulations; Personalized learning; Academic assessment; Technological innovation in medicine; AI tools; AI-assisted medical teaching.

Introduction

Artificial intelligence (AI) is an emerging science that has impacted the educational field, particularly in medical education (1). Its implementation in recent years has grown exponentially, influencing the teaching and learning process (2). Recent studies suggest that the use of AI tools can improve academic performance by providing personalized learning environments (1,3).

Research such as that of Majeed and Sanchez-Gonzalez has compared traditional teaching with AI-based methods, showing that students using AI tools significantly improve their academic performance compared to those following conventional methods (1,4). Taken together, these and other recent studies highlight the potential of AI in terms of academic performance.

Although isolated studies have been conducted on the impact of AI in the performance of medical students, there is no comprehensive synthesis of such impact that considers the benefits, limitations and other considerations of the impact of the use of AI tools on medical student academic performance. This scoping review is relevant as it represents a comprehensive synthesis that allows future researchers to raise hypotheses and questions that will lead to new studies with improved methodologies or directed at topics in which there are currently information gaps. By exploring the current applications, benefits and challenges associated with these technologies, it seeks to provide a comprehensive view of their impact on medical education and contribute to the understanding of how AI innovations are shaping the future of healthcare education. This scoping review aims to identify how the implementation of AI tools influences the academic performance of medical students.

Methods

Eligibility Criteria

In this work (review) are eligible empirical studies (controlled, observational, qualitative and mixed trials), systematic reviews, scoping reviews and case studies available in English and Spanish, published between 2014 and 2024, which seek to examine the impact of the implementation of AI tools in the learning process of medical training, either in teaching, academic performance or clinical simulations. The participants of these studies must be undergraduate or graduate medical students, however, also included are investigations where there is participation of medical educators evaluating the use of AI, providing a complementary perspective on how AI tools affect teaching methodologies and the educational environment in general. Educators' views can enrich our understanding of the role of AI in improving student learning outcomes and adapting curricula to better prepare future medical professionals
Finally, they should evaluate this tool through learning and knowledge acquisition, changes in student motivation and engagement, efficacy in clinical simulations and/or improvement in practical or clinical skills.
Additionally, we have considered specific quality criteria for the studies included in this review. These criteria encompass methodological rigor, sample size adequacy, clarity of outcomes measured, and relevance to the objectives of this review. The eligibility criteria were cross-checked by the authors to ensure a comprehensive assessment of the included research.
The eligibility criteria for this work were cross-checked by the authors of this paper.

Sources of information

The search for information was carried out by reviewing different databases, including PubMed, Scopus and Web of Science, using a systematic search method. This process was divided into three stages: in the first, key words were selected; in the second, filters were applied to reduce the number of articles; and in the third, articles relevant to the topic of study were selected.

Search query

The search queries for this article were created with the help of Chat GPT 3.5. The search terms were selected according to the research question How does the implementation of artificial intelligence tools influence the academic performance of medical students? The same keywords were selected for the three databases, asking ChatGPT 3.5 to create a specific search query for each database consulted (PubMed, Scopus and Web Of Science) with those keywords, combining the lists of terms Artificial Intelligence and performance in medical students, using the AND and OR operators. These search queries were subsequently reconfirmed by 3 reviewers, corresponding to the authors of this scoping review. Examples of the search queries are provided in section A of Multimedia Appendix 1.

Data recording process

Details of the study selection process will be shown by a flow chart (PRISMA-ScR) of the databases searched, included, and excluded. For this scoping review, a table was created for the data record of each article and updated throughout the review process as needed. This table was used to facilitate the selection of relevant information for this review, the elements included in this table are detailed in the following section.

Data elements

The information recorded for this review includes the author(s) of the article, year of publication, country, objective of the study, study design, study participants, data collection, data analysis, and findings.

Categorization of studies

The papers included for this review were classified according to their type of study, being classified as quasi-experimental, randomized controlled trials (RCTs), literature reviews and cross-sectional descriptive studies.

Sources of evidence selection

The PRISMA flow chart in Figure 1 summarizes the selection of studies. The searches were performed on September 13, 2024. The first search was performed in PubMed, where 64 results were initially obtained, after applying the filter by year (2014 to 2024), the number of results was reduced to 58. Subsequently, the search was performed in Scopus, 164 results were

shown, the year filter was applied (2014 to 2024), in addition the filter subject area "medicine" and type of document "article" was selected, obtaining 54 results. The last database searched was Web of Science, the generic result was 87 documents; once the filter year (2014 to 2024) and type of document "article" was applied, 61 results were obtained.

Finally, after the search query, a total of 173 studies containing terms Artificial Intelligence and academic performance in medical students were found. Those results were aggregated in the Rayyan platform, for the detection of duplicate articles and to facilitate the process of screening relevant articles for the present scoping review. Rayyan detects duplicates by analyzing various factors such as title, authors, journal, and year of publication. It assigns a similarity percentage to potential duplicates, allowing reviewers to identify and resolve them efficiently. The platform also offers an Auto-Resolver feature that can automatically handle duplicates with high similarity scores (95% or higher), streamlining the deduplication process while maintaining accuracy.

After eliminating duplicate studies, 108 articles were identified for review. The articles found from the search strategy were independently evaluated by three reviewers, corresponding to the authors of the present review. Reading the title and abstract of each article reduced the total number to 21, categorized in Rayyan as "Included" and "Maybe". The remaining 87 articles were excluded under 3 criteria, with no information on students' academic performance (21), learning methodologies without the use of AI (53), and academic performance according to learning approaches (13).

Subsequently, a complete reading of those 19 articles was carried out, approving 8 documents.

To resolve discrepancies among reviewers during the evaluation process, a consensus methodology was employed. When disagreements arose regarding the inclusion or exclusion of specific articles, discussions were held among the reviewers to reach a mutual agreement based on predefined criteria and the relevance of each study to the review's objectives.

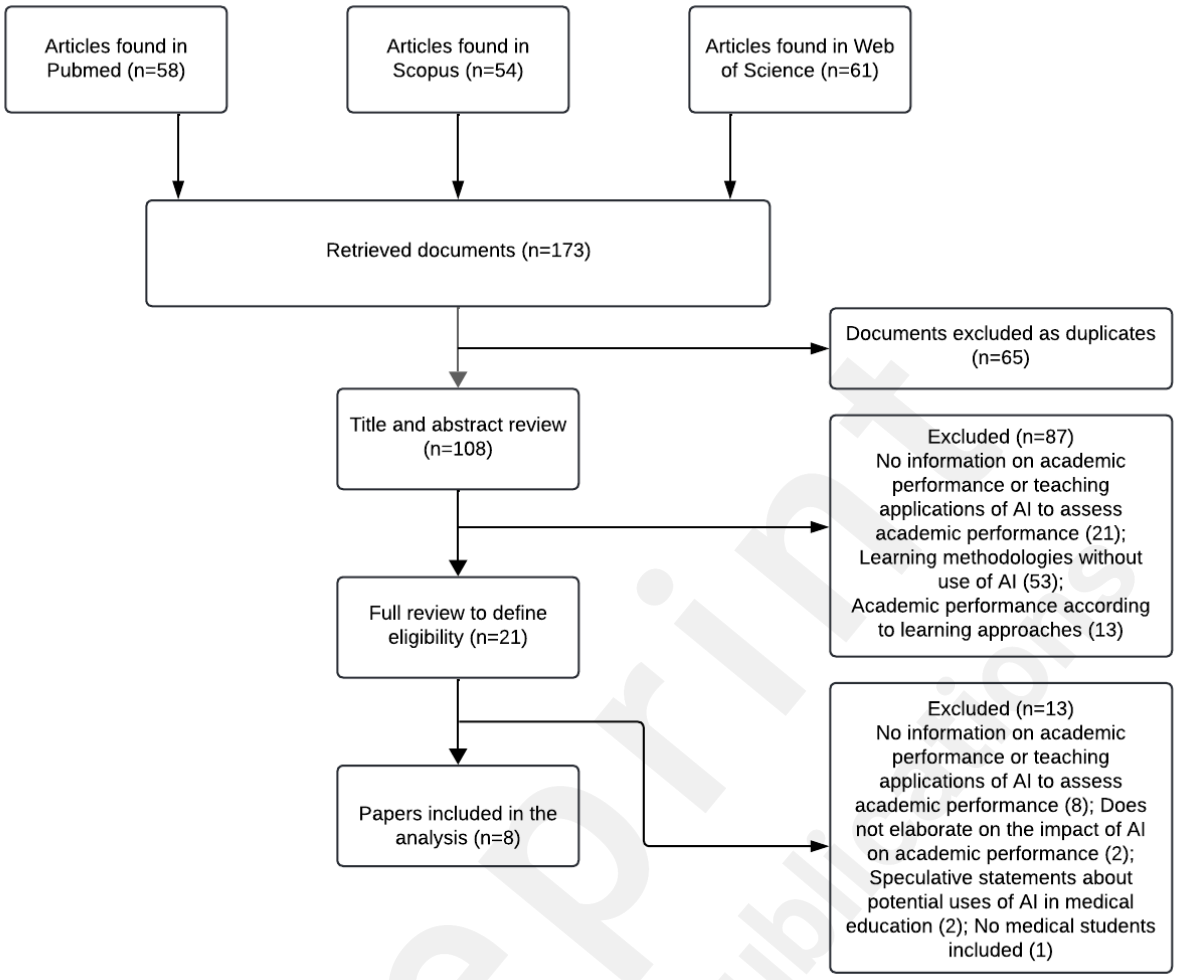

Figure 1. Flow diagram of the studies included in this review.

Results

This review includes the analysis of 8 studies selected for review according to the exclusion criteria. The main results of this review are listed in Tables 1 to 2. In Table 1, 4 articles comparing traditional teaching methods and those incorporating AI were selected. In Table 2, 4 articles analyzing the impact of AI on simulation and personalized learning were selected.

Table 1. Summary of studies comparing traditional teaching methods and those incorporating AI.

Author (Year, Country)	*Objective of the study*	*Study design*	*Participants*	*Collection of information*	*Data analysis*	*Findings*
Ayesha Isani Majeed (2024, Pakistan) (4)	To compare AI-based teaching with the traditional chest X-ray approach	Randomized controlled trial	40 radiology residents. Group A (traditional teaching; n=20) and Group B (AI-based	Pre-intervention assessments and diagnostic ability assessment questionnaires.	SPSS v25; independent-sample and chi-square tests.	Group A in the post-test obtained an average score of 9.40 ± 0.59. Group B obtained an average post-

	to detect COVID-19.		teaching; n=20).			test score of 7.75 ± 1.12. Significant difference (*P* < .001).
Sanchez-Gonzalez M, Terrell M (2023, United States) (1)	Evaluating the effectiveness of voice-over PowerPoint lessons and artificial intelligence technology to facilitate learning	Randomized controlled trial	California NorthState University first-year medical students with no prior knowledge of basic science.	Multiple choice questionnaire and surveys	Student's t test and Cohen's d calculation.	Students using PPT+AI scored significantly higher on questions: 75.2 ± 3.3% versus 68.3 ± 2.9% in the traditional PPT group (*P* < .043, d = 0.54).
Nofal (2024, Jordan) (5)	To evaluate medical students' use of general surgery resources and relate them to their grades.	Retrospective descriptive study by means of a survey.	Jordanian medical students from six universities.	7-item survey conducted between March and June 2023.	SPSS version 26, descriptive statistics, chi-square test.	Traditional face-to-face education was the most popular and efficient method of study in general surgery, obtaining the best grades. Most students did not receive formal education on the use of AI during their medical training.
Gan W, et al. (2024, China) (6)	To evaluate the accuracy of ChatGPT in solving multiple choice questions	Randomized controlled trial	129 third-year medical students of Jinan University who have passed the courses of anatomy,	Multiple-choice tests, final exam scores, ChatGPT-4.0 answers to multiple-choice questions and	SPSS software, chi-square, Kolmogorov-Smirnov test, Mann Whitne	The group using ChatGPT answered more orthopedic questions correctly than the control

	on orthopedics and also its short-term effects as a tool for learning.		physiology, pathophysiology, pathology, biochemistry and diagnostics.	telephone interviews.	y U test.	group. In addition, they performed significantly better on their OB/GYN and surgical final exams. ChatGPT-4.0 answered 70.6% of the orthopedic multiple-choice questions correctly. ChatGPT-4.0 is a good teaching aid for basic medical knowledge, but not for diagnostic and treatment tasks.

Table 2. Summary of studies analyzing the impact of AI on simulation and personalized learning.

Author (Year, Country)	*Objective of the study*	*Study design*	*Participants*	*Collection of information*	*Data analysis*	*Findings*
Castellano, et al. (2024, Chile) (7)	To evaluate the impact of a gamified platform in anatomy teaching, with a focus on personalization of learning to	Quasi-experimental study	131 medical students from two universities. Participants were divided into control and experimental groups.	Questionnaire, data provided by the application and surveys.	Analysis quantitative and qualitative; t-test.	In both universities the experimental group obtained better results in the post-test, with an average increase of 4-5 points in the post-test (8/15 to 12-13/15), and

	improve the academic performance of medical students.					a significant difference (*P*=.041).
Yilmaz R., et al.. (2024, Canada) (8)	To explore optimal feedback methodologies to improve simulated bimanual surgical skills in brain tumor resection.	Randomized controlled trial.	120 first- to fourth-year medical students from four universities in the province of Quebec. They were randomly assigned to four different feedback groups	Data were collected in 60-minute simulation sessions without follow-up.	Parametric statistical tests, equivalent nonparametric statistical analysis. Statistical analysis was performed using IBM SPSS Statistics, version 27.	Group 1 (no feedback): final average of 7.30 +- 1.69. Group 2 (numerical feedback): 8.40 +- 2.06. Group 3 (visual feedback): 9.10 +- 1.82. Group 4 (visuospatial feedback): 9.19 +- 1.66. There are significant differences in benchmarks achieved between the feedback and non-feedback groups.
Cabañuz, Clara, García-García, Mar (2024, Spain) (3)	To explore the use of AI using ChatGPT for test-taking exams in medical students.	Descriptive study	228 students attended the exam for the January-February session of the 2022-2023 anatomopathological diagnostic and therapeutic procedure	Results of the exam, excluding practical questions and questions appealed by students.	Quantitative methods and statistical descriptors (median and mean) and Student's t-test.	The mean course grade was 5.35 and the median was 5.29. While ChatGPT answered 47 of 68 questions correctly, giving a 10-weighted score of 6.14, above the course mean and

			s course.			median.
Garabet R., et al. (2023, United States) (9)	Determine AI performance on the question sets, extending the initial findings of GPT-3 and GPT-3.5 using the newer model (GPT-4), reporting performance by subject and discipline.	Cross-sectional descriptive study	ChatGPT	The answers were evaluated as correct or incorrect by comparing them with the comments provided after selecting the answer option in the AMBOSS question bank.	IBM SPSS version 28. Unpaired chi-square tests.	ChatGPT accurately answered 86% of a total of 1300 questions. In exams such as the USMLE STEP 1, ChatGPT-4 showed proficiency in all tested subjects. ChatGPT could be a good addition to medical school curricula.

Summary of results

In general, there is a tendency in the studies to report better academic performance in those students who have used AI in their learning compared to those who use methods that do not include AI and to highlight the ability of extensive language models such as ChatGPT to correctly answer questions of alternatives related to medical subjects, in addition to being useful as assistance in the elaboration of these. In addition, it is highlighted that the use of AI favors better learning and performance in practical skills.

Before dividing the included articles into categories, they were examined in a general way to determine the distribution of articles considering the date of publication and type of study. The included articles correspond to the period from 2022 to 2024. Although, the search was conducted from 2014 onwards, only in the last three years the studies that were reviewed relate the use of artificial intelligence with the academic performance of medical students, highlighting that most of these belong to the year 2024. Four randomized controlled trials, three descriptive studies, one literature review and one quasi-experimental study were identified.

Analysis of natural categories

Following the general analysis, the articles were read again in full to identify common themes on the basis of which they could be grouped, 2 main general themes related to the impact of artificial intelligence on academic performance were identified. These were the comparison of teaching methods and the analysis of personalized learning using AI and its impact on

simulations. The studies were then grouped into natural categories according to the type of theme they possessed. The following natural groups were identified: 1) Comparison of traditional teaching methods and those incorporating AI 2) Analysis of the impact of AI on simulation and personalized learning. In the first group, 4 articles were selected, the general objective of which is to compare the effectiveness of traditional teaching methods with those that integrate artificial intelligence tools, evaluating their impact on academic performance and the learning experience of students, in order to identify opportunities to optimize the use of AI in medical education. In the second group, 4 articles were selected, being the general objective to analyze the impact of AI in simulation and personalized learning in medical education, evaluating its potential in the acquisition of knowledge, clinical skills and adaptability to the individual needs of students, through tools such as gamified platforms, simulations and large language models such as ChatGPT.

Discussion

General discussion

This review provides a general overview of the topic, categorizing the literature reviewed according to its main general themes. In each category, common concepts were explored and potential future applications of AI in medical education were identified to enhance the learning and academic performance of medical students. It is observed in each of the categories that AI has great potential for development to support the learning of medical students, improving their academic performance.

The impact of artificial intelligence tools on the academic performance of medical students is still a very recent and growing topic, whose study has become more frequent in recent years, focusing mainly on the teaching of basic sciences and knowledge that have a strong visual component, such as anatomy or the interpretation of radiographs, but also exploring their uses in clinical simulations and surgical scenarios.

Discussion of studies that compare traditional teaching methods and those that incorporate AI.

The main objective of these studies was to compare the use of AI as part of the teaching method with other traditional teaching methods or non-AI learning resources, such as YouTube channels and medical websites, relating the use of AI to the academic performance of students in a test specifically designed to assess the knowledge taught during the study, or in their grades obtained in final exams in certain subjects. The most commonly used study method was the randomized controlled trial(1,4,6), using control groups that had exactly the same number of participants as the experimental groups (1,4) or a similar number (6). Study participants were grouped by academic year. Only one study used a retrospective design, collecting information by means of a survey (5).

In all randomized controlled trials, the group that received AI education had significantly better results in the tests performed, either in the total score obtained in the test or in those questions of greater difficulty (1,4,6).

Only one study concluded that traditional face-to-face teaching through tutor lessons and surgery books is the best approach to achieve higher academic performance, being found above the other methods, however it highlights the low sample of students using artificial intelligence tools in their study (27.2%), the lack of information regarding the frequency of use of these tools and regarding which AI tools are used by students. It is also important to note that 72.8% of the students in this study did not receive formal education on the use of AI during their training, being concordant with the large proportion of medical students reporting the same (5,10). There is a general trend in the studies to consider an early incorporation of AI in teaching and learning, noting that the specific impact of AI in education is still largely unexplored (11) and that the lack of professional training and critical evaluation of AI technologies are an impediment to the next generation of physicians (12). Finally, gan et al. point out that the incorporation of tools such as AI by students into their learning styles is inevitable and that society should have high standards for teaching methods, teaching content, and student assessment (6).

Studies highlight the use of AI in student assessment and the lower time investment it requires compared to traditional methods (1,2,4,13). All studies highlight the ability of AI to serve as a learning support (4-6,13), although some extensive language models such as ChatGPT-4.0 may not be as good aids for tasks such as diagnosis and treatment, but they are good for teaching basic medical knowledge, thanks to their ability to answer questions instantaneously and collect and collate data quickly (6).

The findings of the articles reviewed are in agreement with the findings of older studies addressing the comparison between AI and traditional teaching methods, documenting favorable results for AI (14,15), and more recent studies document that AI is able to support the learning of medical subjects, such as anatomy, physiology, and biochemistry (16,17), and to create engaging educational material that increases student engagement and motivation for learning (18,19).

Finally, the effectiveness of AI as part of educational strategies to improve academic performance has been highlighted in recent times (13). It has been observed how AI is able to reduce the time between learning and assessment, thanks to its ability to provide instant feedback available at any time (1,20). According to other studies, this feedback is able to help students identify weak areas for improvement (21,22) and improve their overall performance (21).

Discussion of studies analyzing the impact of AI on simulation and personalized learning.

The main objective of these studies was to analyze the impact of AI on simulation and personalization of learning in medical education. Among the approaches investigated, the use of gamified platforms to improve anatomy teaching, this platform included a virtual assistant and personalized systems to adapt learning to the needs of medical students (7); the evaluation of different types of feedback in simulated surgical skills (8); and the ability of ChatGPT to solve test-type questions and generate new questions (3,9).

One of the methodologies used was a quasi-experimental study where medical students were divided into control and experimental groups (7). Another was a randomized controlled trial that evaluated students in surgical simulations using the NeuroVR platform (provides a virtual environment with real-time feedback) students received different types of feedback

(numerical, visual, visuo-spatial and no feedback) (8). In addition, two studies performed descriptive analyses, one compared ChatGPT responses with those of students on test-type questions and further evaluated how ChatGPT could generate new questions (3), and another study used questions from different medical disciplines from the AMBOSS educational platform to evaluate the performance of ChatGPT-4, comparing the responses delivered by the AI with that of medical students (9).

As for the results, the study evaluating the use of gamified platforms showed a significant improvement in the academic performance of students who used this tool, with an average increase of 4-5 points in the post-test of the experimental group (8/15 to 12-13/15), in addition 85-88% of students indicated that the use of gamified platforms was useful for learning (7). Moreover, the use of the NeuroVR platform visual and visuo-spatial feedback was the most effective in learning surgical skills, compared to those who did not receive feedback or only received numerical feedback. Finally, one of the studies using ChatGPT, the AI answered 69% of the test-type questions correctly, outperforming the average student performance of 63.5%. The other study using ChatGPT-4.0 showed an overall accuracy of 86% out of a total of 1300 questions, compared to medical students ChatGPT-4.0 outperformed in several disciplines, for example, in the cardiovascular system, the model achieved an accuracy of 90%, while students scored an average of 64%. Similarly, in the musculoskeletal system, ChatGPT achieved an accuracy of 94%, in contrast to 61% for students (9). Some studies suggest that the accuracy with which ChatGPT answers a question is related to both the complexity of that question and its ability to associate it with the information contained in its data corpus (23,24).

Taken together, the studies reviewed demonstrate that artificial intelligence has a significant impact on medical education, whether through simulation platforms, gamification, GPT Chat, among others. The ability of AI to provide adaptive feedback and generate personalized scenarios could increase academic performance, as evidenced in areas such as anatomy, surgical skills and test-like question solving. Along these lines, AI-powered platforms are able to analyze students' strengths and weaknesses, providing personalized learning experiences (16,21) that can be of great benefit to the learner. Another impact on teaching is simulation, which offers increasingly realistic and accurate tools, allowing students to practice in a controlled and safe environment. Although interaction with real patients is still essential, the prior use of these simulated tools would help to minimize risks (25,26). On the other hand, among the most recent tools, ChatGPT stands out, since its introduction, it has shown potential in the area of medicine, some studies have proposed that the use of ChatGPT facilitates personalized and interactive learning, which can improve evaluation processes and generate a continuous feedback loop (24,27,28). However, it also highlights the need to develop ethical frameworks to ensure the responsible use of AI, specifically ChatGPT considering additional risks related to the accuracy and validity of the information generated, ethics in data handling, and information privacy and security. As obtained in the results of the articles, some ChatGPT responses may be incorrect, which may raise doubts about their reliability, especially if they are used by people without the ability to detect or correct errors (3,29). Therefore, AI in medical education should be used as a support tool, consciously and always complemented with human critical judgment.

Conclusions

Conclusions on studies comparing traditional teaching methods and those incorporating AI:

Overall, most of the studies reviewed demonstrate that methodologies using AI can significantly improve the academic performance of medical students compared to traditional teaching methodologies. The use of AI in medical education has attractive advantages such as reduced time investment and its ability to provide instant feedback, which contributes to improved academic performance. Thus, the incorporation of AI tools as part of teaching strategies by educational institutions for medical students could improve their academic performance and learning during their academic training, but more research is still needed in this field to determine those areas of knowledge in which AI could have a greater impact on the performance of medical students and those in which it should be adjusted.

There is a trend to consider an increase in the incorporation of AI in education, but there are still many students who have not received formal education during their training on the use of these tools, which could limit the use of these tools and be an impediment to the training of the next generation of physicians.

Conclusions on studies analyzing the impact of AI on simulation and personalized learning:

AI has significant potential to transform medical education, especially in the area of simulation. The studies reviewed indicate that gamified platforms and simulators have been shown to markedly improve students' academic performance and practical skills. In addition, the potential of AI to enrich learning in various medical specialties is highlighted. However, it is essential that the development and implementation of these technologies be done in a careful and informed manner to maximize their benefits.
Artificial intelligence is not only revolutionizing medical education through simulations and gamification, but also offers opportunities to personalize learning and improve academic performance. However, it is vital to take a critical and evidence-based approach to ensure that these innovations are effectively implemented and continue to benefit students in their professional training. Further research is essential to identify additional predictors of effective learning and improve assessment methodologies.

Strategic recommendations and research proposals for the integration of AI in medical education.

To enhance the integration of artificial intelligence in medical education, it is essential that educators receive specific training in the use of these tools, incorporating them as part of their continuing professional development. Likewise, the delivery of knowledge on the practical application of AI in diagnosis, treatment and medical learning should be included in undergraduate and postgraduate academic programs. As for future lines of research, it is a priority to evaluate the impact of AI in specific areas of medical knowledge, identifying those with the greatest potential to improve student performance. In addition, it is essential to develop more precise evaluation methodologies to measure the benefits of the personalization of learning through AI and to explore the ethical implications of its incorporation into medical training.
Reliance on algorithms and automated systems can perpetuate biases, affect data privacy, and reduce human clinical judgment. Therefore, it is crucial that educational institutions include ethical training in their curricula, preparing future professionals to meet these challenges,

ensuring that their application equitably benefits all patients and communities.
Finally, it is crucial to recognize that the development and implementation of these technologies require significant investments, which can be a challenge for universities in countries with limited resources. To ensure that all students have access to these advanced educational tools, it is essential to foster collaborations between institutions, governments and the private sector, as well as to explore sustainable funding models. Only in this way is it possible to ensure that artificial intelligence will not become an exclusive privilege, but an accessible resource that enriches medical education and improves healthcare globally.

Limitations

The analysis of this topic was limited to 3 major databases and important articles from other databases could have been ignored. In addition, another limitation is the recent incorporation of studies dealing with artificial intelligence in medical education, which has resulted in a small number of investigations that meet the inclusion criteria. Although the available literature provides relevant information, the limited number of studies makes it difficult to generalize the results, highlighting the need for further research on the use of AI.

Conflicts of Interest

None declared

Abbreviations

AI: Artificial intelligence
JMIR: Journal of Medical Internet Research
RCT: randomized controlled trial

Multimedia Appendix 1

Supplementary material Multimedia Appendix 1

References

1. Sanchez-Gonzalez M, Terrell M. Flipped Classroom With Artificial Intelligence: Educational Effectiveness of Combining Voice-Over Presentations and AI. Cureus. 2023 Nov;15(11):e48354.
2. Chan KS, Zary N. Applications and Challenges of Implementing Artificial Intelligence in Medical Education: Integrative Review. JMIR Med Educ. 2019 Jun 15;5(1):e13930.
3. Cabañuz C, García-García M. [ChatGPT is an above-average student at the Faculty of Medicine of the University of Zaragoza and an excellent collaborator in the development of teaching materials]. Rev Espanola Patol Publicacion Of Soc Espanola Anat Patol Soc Espanola Citol. 2024 Jun;57(2):91–6.
4. Majeed AI. Comparison of Artificial Intelligence-based learning with the traditional method in the diagnosis of COVID-19 chest radiographs among postgraduate radiology residents. KHYBER Med Univ J. 2024 Jun 30;16(2):140–4.
5. Nofal MN, Al Awayshish MM, Yousef AJ, Alamaren AM, Al-Rabadi ZI, Haddad DS, et al. General surgery educational resources for Jordanian medical students. Surg Open Sci. 2024 Aug;20:62–5.

6. Gan W, Ouyang J, Li H, Xue Z, Zhang Y, Dong Q, et al. Integrating ChatGPT in Orthopedic Education for Medical Undergraduates: Randomized Controlled Trial. J Med Internet Res. 2024 Aug 20;26:e57037.
7. Castellano MS, Contreras-McKay I, Neyem A, Farfán E, Inzunza O, Ottone NE, et al. Empowering human anatomy education through gamification and artificial intelligence: An innovative approach to knowledge appropriation. Clin Anat N Y N. 2024 Jan;37(1):12–24.
8. Yilmaz R, Fazlollahi AM, Winkler-Schwartz A, Wang A, Makhani HH, Alsayegh A, et al. Effect of Feedback Modality on Simulated Surgical Skills Learning Using Automated Educational Systems- A Four-Arm Randomized Control Trial. J Surg Educ. 2024 Feb;81(2):275–87.
9. Garabet R, Mackey BP, Cross J, Weingarten M. ChatGPT-4 Performance on USMLE Step 1 Style Questions and Its Implications for Medical Education: A Comparative Study Across Systems and Disciplines. Med Sci Educ. 2024 Feb;34(1):145–52.
10.Pucchio A, Rathagirishnan R, Caton N, Gariscsak PJ, Del Papa J, Nabhen JJ, et al. Exploration of exposure to artificial intelligence in undergraduate medical education: a Canadian cross-sectional mixed-methods study. BMC Med Educ. 2022 Nov 28;22(1):815.
11.Xie H, Hwang GJ, Wong TL. Editorial Note: From Conventional AI to Modern AI in Education: Re- examining AI and Analytic Techniques for Teaching and Learning. Educ Technol Soc. 2021;24(3):85-88.
12.Pucchio A, Eisenhauer EA, Moraes FY. Medical students need artificial intelligence and machine learning training. Nat Biotechnol. 2021 Mar;39(3):388–9.
13.Varma JR, Fernando S, Ting BY, Aamir S, Sivaprakasam R. The Global Use of Artificial Intelligence in the Undergraduate Medical Curriculum: A Systematic Review. Cureus. 2023 May;15(5):e39701.
14.Gorby GL. Use of Verbot technology to enhance classroom lecture. Acad Med J Assoc Am Med Coll. 2001 May;76(5):552–3.
15.Michael J, Rovick A, Glass M, Zhou Y, Evens M. Learning from a Computer Tutor with Natural Language Capabilities. Interact Learn Environ. 2003 Dec 1;11(3):233–62.
16.Temsah MH, Alhuzaimi AN, Almansour M, Aljamaan F, Alhasan K, Batarfi MA, et al. Art or Artifact: Evaluating the Accuracy, Appeal, and Educational Value of AI-Generated Imagery in DALL·E 3 for Illustrating Congenital Heart Diseases. J Med Syst. 2024 May 23;48(1):54.
17.Luke WANV, Seow Chong L, Ban KH, Wong AH, Zhi Xiong C, Shuh Shing L, et al. Is ChatGPT “ready” to be a learning tool for medical undergraduates and will it perform equally in different subjects? Comparative study of ChatGPT performance in tutorial and case-based learning questions in physiology and biochemistry. Med Teach. 2024 Nov;46(11):1441–7.
18.Sauder M, Tritsch T, Rajput V, Schwartz G, Shoja MM. Exploring Generative Artificial Intelligence-Assisted Medical Education: Assessing Case-Based Learning for Medical Students. Cureus. 2024 Jan;16(1):e51961.
19.Safranek CW, Sidamon-Eristoff AE, Gilson A, Chartash D. The Role of Large Language Models in Medical Education: Applications and Implications. JMIR Med Educ. 2023 Aug 14;9:e50945.
20.Siemens G, Long P. Penetrating the Fog: Analytics in Learning and Education. Educ Rev. 2011 Jan 1;5:30–2.
21.Karabacak M, Ozkara BB, Margetis K, Wintermark M, Bisdas S. The Advent of Generative Language Models in Medical Education. JMIR Med Educ. 2023 Jun 6;9:e48163.
22.Hooda M, Rana C, Dahiya O, Rizwan A, Hossain MS. Artificial Intelligence for Assessment and Feedback to Enhance Student Success in Higher Education. Math Probl Eng. 2022 Jan 1;2022(1):5215722.

23.Carrasco J, García E, Sánchez D, Porter E, Puente L, Navarro J, et al. ¿Es capaz “ChatGPT” de aprobar el examen MIR de 2022? Implicaciones de la inteligencia artificial en la educación médica en España. ChatGPT Capab Passing 2022 MIR ExamImplications Artif Intell Med Educ InSpain. 2023 Feb 16;4(1).
24.Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment. JMIR Med Educ. 2023 Feb 8;9:e45312.
25.Doumouras AG, Engels PT. Early crisis nontechnical skill teaching in residency leads to long-term skill retention and improved performance during crises: A prospective, nonrandomized controlled study. Surgery. 2017 Jul;162(1):174–81.
26.Haenlein M, Kaplan A. A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. Calif Manage Rev. 2019 Aug 1;61(4):5–14.
27.Breeding T, Martinez B, Patel H, Nasef H, Arif H, Nakayama D, et al. The Utilization of ChatGPT in Reshaping Future Medical Education and Learning Perspectives: A Curse or a Blessing? Am Surg. 2024 Apr;90(4):560–6.
28.Hosseini M, Gao CA, Liebovitz DM, Carvalho AM, Ahmad FS, Luo Y, et al. An exploratory survey about using ChatGPT in education, healthcare, and research. PloS One. 2023;18(10):e0292216.
29.Zhuo TY, Huang Y, Chen C, Xing Z. Exploring AI Ethics of ChatGPT: A Diagnostic Analysis. ArXiv. 2023;abs/2301.12867.

Supplementary Files

Multimedia Appendixes

Search Query.
URL: http://asset.jmir.pub/assets/7f3cd0a1cdbe02f4ea41475bbeaa307d.docx