ACCSCIENCE PUBLISHING

Artificial Intelligence in Health

REVIEW ARTICLE

The role of artificial intelligence in higher medical education and the ethical challenges of its implementation

Mark Perkins[1,2†] **and Agnieszka Pregowska**[3†*]

[1]Collegium Prometricum, The Business School for Healthcare, Sopot, Poland

[2]Royal Society of Arts, London, United Kingdom

[3]Department of Information and Computational Science, Institute of Fundamental Technological Research, Polish Academy of Sciences, Warsaw, Poland

Abstract

Artificial intelligence (AI) is penetrating higher medical education; however, its adoption remains low. A PRISMA-S search of the Web of Science database from 2020 to 2024, utilizing the search terms "artificial intelligence," "medicine," "education," and "ethics," reveals this trend. Four key areas of AI application in medical education are examined for their potential benefits: Educational support (such as personalized distance education), radiology (diagnostics), virtual reality (VR) (visualization and simulations), and generative text engines (GenText), such as ChatGPT (from the production of notes to syllabus design). However, significant ethical risks accompany AI adoption, and specific concerns are linked to each of these four areas. While AI is recognized as an important support tool in medical education, its slow integration hampers learning and diminishes student motivation, as evidenced by the challenges in implementing VR. In radiology, data-intensive training is hindered by poor connectivity, particularly affecting learners in developing countries. Ethical risks, such as bias in datasets (whether intentional or unintentional), need to be highlighted within educational programs. Students must be informed of the possible motivation behind the introduction of social and political bias in datasets, as well as the profit motive. Finally, the ethical risks accompanying the use of GenText are discussed, ranging from student reliance on instant text generation for assignments, which can hinder the development of critical thinking skills, to the potential danger of relying on AI-generated learning and treatment plans without sufficient human moderation.

Keywords: Artificial intelligence; Metaverse; Medical education; Education system; Ethics

1. Introduction

Medical practice, which heavily relies on advancements in medical education, is one of the fastest-moving fields, frequently testing technological innovations through pilot trials and proof-of-concept studies.[1] Artificial intelligence (AI) now stands at the forefront of these innovations, offering many benefits, such as effective tools for analyzing and processing large datasets quickly – tasks that would be impossible for humans to accomplish.

†These authors contributed equally to this work.

***Corresponding author:** Agnieszka Pregowska (aprego@ippt.pan.pl)

Citation: Perkins M, Pregowska A. The role of artificial intelligence in higher medical education and the ethical challenges of its implementation. *Artif Intell Health*. 2025;2(1):1-13. doi: 10.36922/aih.3276

Received: March 26, 2024

Revised: April 29, 2024

Accepted: July 1, 2024

Published Online: October 21, 2024

One important area in healthcare involves electronic health records, which can serve as input data for AI and be processed quickly. However, such datasets not only contain sensitive content but also constitute ethical risks, especially when data collection is subject to various forms of bias[2] and is exposed to a large number of hostile attacks.[3] More concerning is that most medical researchers treat AI as a black box, leaving its ethical risks concealed.[4] A strong argument can be made that the successful application of AI in medical practice will depend on addressing legitimate concerns about misunderstandings of its principles and data risks, in accordance with evolving bioethical principles.[5] In a field such as medicine, which is critically related to issues of life and health,[6] it is particularly important to explain and address the impact of AI on its essence and principles in medical educational programs, both in terms of how it works and its underlying ethical assumptions. For medical practitioners to use AI-based solutions effectively in their work, they must first learn how to use them correctly during their training.

Moreover, AI-based solutions may be more vulnerable to attacks compared to other approaches, such as statistical methods.[7] It is also worth stressing that, especially in the field of medicine, deep neural networks with many layers (such as highly complex architectures) are commonly applied. This may contribute to AI models being more susceptible to overfitting, where the neural network memorizes the training data rather than generalizing from it. In this context, statistical methods are composed of simpler models with fewer parameters, which may lead to easier interpretation of the model.[8]

A significant limitation of AI is its dependence on data.[9] In particular, the essence of AI, comprising algorithms for learning complex patterns and making accurate predictions, has a core sensitivity feature: the quality and representativeness of the training data. Inaccuracies in the training data significantly affect the efficiency and accuracy of the results obtained, potentially skewing outcomes and leading to ethical consequences that oppose the institution's goal. Indeed, it can be said that the quality and output of AI algorithms are directly dependent on the medical data used to develop, test, and validate them. Therefore, a key issue in using AI in medicine is the reliability of biomedical data obtained from patients, which must be compiled and categorized in an ethical manner. Unlike AI-based models, statistical methods can work with smaller datasets, and the optimal selection of data may help minimize data errors more efficiently. The heavy data dependence on AI-based solutions also makes them vulnerable to developing learning patterns based on biased and faulty training data. If the input data is not representative of the real-world population or reflects historical biases and inequalities, AI can learn and perpetuate these biases. For example, a language model trained on text from certain online communities may accidentally learn and replicate the biases expressed in that community. A lack of diversity in the ethical standpoint of AI researchers may also contribute to bias issues. Moreover, the algorithms themselves may introduce or amplify algorithmic errors due to their inherent operational principles.

Another challenge related to data is security. Compared to traditional statistical methods, AI-based algorithms are more susceptible to adversarial attacks that exploit security vulnerabilities, such as sensitivity to even low noise in the input data.[10] Traditional methods are more deterministic, making them more resistant to such attacks.

In this paper, we analyze the technical and ethical risks associated with certain AI applications in medical education, exploring the potential benefits and risks of these technologies in practice, the awareness of students and practitioners regarding these issues, and the latest scientific research in this area.

2. An overview of current research activity in AI, medical education, and ethics

In this paper, we conducted a systematic review of research on AI, medical education, and ethics based on the PRISMA academic review process and its extensions, including PRISMA-S.[11] Resources written in English from the Web of Science (WoS) database were considered, excluding PhD theses and any material not related to AI or education. Our searches for the terms "AI," "education," and "medicine" yielded 488 resources, of which 34 addressed ethical issues. Figure 1 presents the participation rate in % of individual areas of the world in research relating to AI in medical education (Figure 1A) and AI in medical education, taking into account ethical issues (Figure 1B). These results highlight both the very low participation of low-income countries in research and a lack of focus on ethics. However, the study also included searches involving the search terms "artificial intelligence," "medicine," and "ethics" (AI+med+ethics), which yielded 328 results, giving a higher result when education as a whole is considered. The sources included were selected to answer the research question, "What multi-criteria impact will AI have on higher education in the field of medicine?" First, duplicate records in the database were excluded. In the second step, records whose titles and abstracts were not related to the subject of the analysis were excluded. Then, records that were not accessible were disabled. In the final stage of the search, records without information concerning the topic of consideration were excluded from the analysis. Finally,

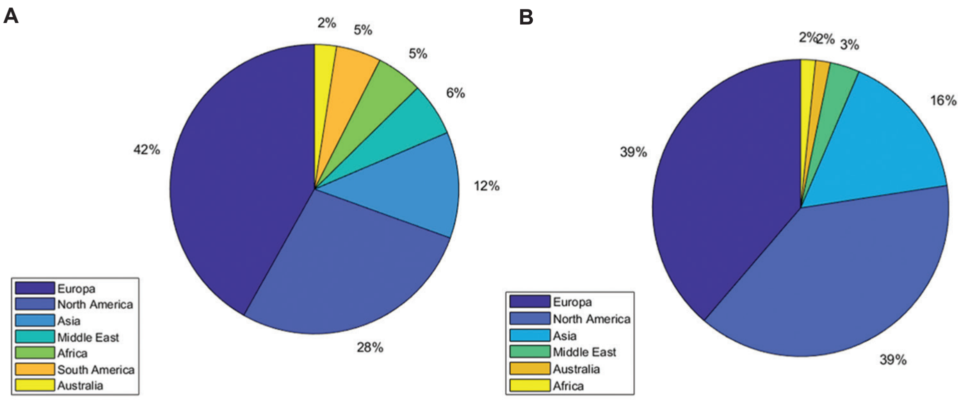

Figure 1. Geographical distribution of papers related to artificial intelligence (AI) in medical (med) education (ed) on the Web of Science. (A) Papers including the terms AI+med. (B) Papers including the terms AI+med.

94 documents were taken into account. This investigation has two limitations. First, the study only takes into account the WoS database, which is the most restricted of its type (although this ensures the integrity of the dataset), and second, only publications written in English were included in the systematic review, which may cause a potential language bias.

The results of 328/488 resources may seem low. This view is supported by Lee *et al.*,[12] who noted that AI is a relatively new concept in medical education. More recently, as a result of an exhaustive search in four databases (PubMed, Embase, Scopus, and WoS) during the period 2020–2024, Weidener and Fischer[13] affirmed that there is "a scarcity of literature on teaching AI ethics in medical education, with most of the available literature being recent and theoretical" (ibid., p. 399). The study shows that the major studies (about 90%) in the field of AI ethics were published in the years 2020–2024, which coincides with the dynamic development of AI. This is largely due to the fact that currently solutions based on AI can be implemented in practice, and there is a need to consider all risks, both ethical and practical (technical). Since we analyze the status of development and implementation of the general guidance on the ethics of AI in the field of medical education (with special emphasis on practical implications), in this study, we concentrate on the time frame in which the most dynamic development of the field of AI ethics occurs. In addition, the analysis highlighted a research gap in low-income countries. One of the reasons may be the lack of access to the latest technologies, which often involves significant costs. The lack of research in this area also translates into a potentially low level of implementation of AI in practice. Indeed, it is evident that the results of AI+med+ed are a small proportion of those for AI+med, and that the results for AI+med+ed+ethics are an even smaller proportion of AI+med (Figures 2A and 2B). Overall, there is an underrepresentation of research in the developing world, despite the recognized importance of AI+med+ed+ethics.

A similar situation was found when a search for the terms AI+med+radiology and AI+med+XR was conducted (Figure 3A and B).

Almost half of the occurrences of AI+med+radiology were found in North America (49%), compared to only 7% in Europe. On the other hand, almost the opposite was found regarding AI+med+XR: North America (47%) and Europe (19%). This suggests that research and awareness of AI and radiology are more advanced in North America than in Europe and that the opposite is true concerning AI and XR. Asian results were similar in both cases (25% and 21%), but as a large developing territory, Africa was substantially underrepresented (2% and 3%).

3. AI in medical education-some practical applications

AI is increasingly seen as a significant resource for medical education that will permeate all areas and become integral. AI is being applied in several different types of medical fields, including technical support and distance learning, data analysis and interpretation, 3D modeling and remote virtual surgery, and text production by AI-powered text generation engines (GenText) such as ChatGPT (Chat Generative Pre-trained Transformer). A study by Civaner *et al.*[14] showed that 80% of medical students perceive AI as a technology supporting both the process of education and health care, although the study also revealed concerns among the medical community about AI undermining their skills and negatively impacting the patient-doctor relationship (50% and 40% of respondents, respectively). These concerns are not shared by biomedical physicians, more than 80% of whom see AI as a support tool, not a risk

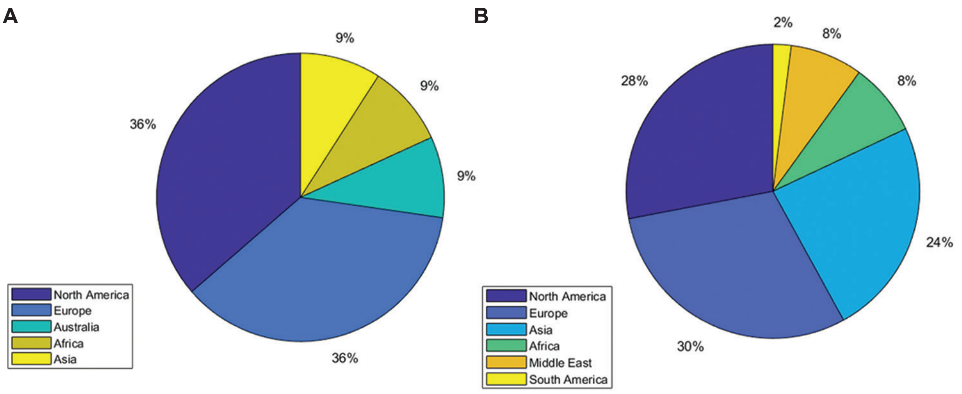

Figure 2. Geographical distribution of papers related to artificial intelligence (AI) in medical (med) education (ed) and ethics (ethics) based on the Web of Science. (A) Distribution of papers on AI+med+ed. (B) Distribution of papers on AI+med+ed+ethics.

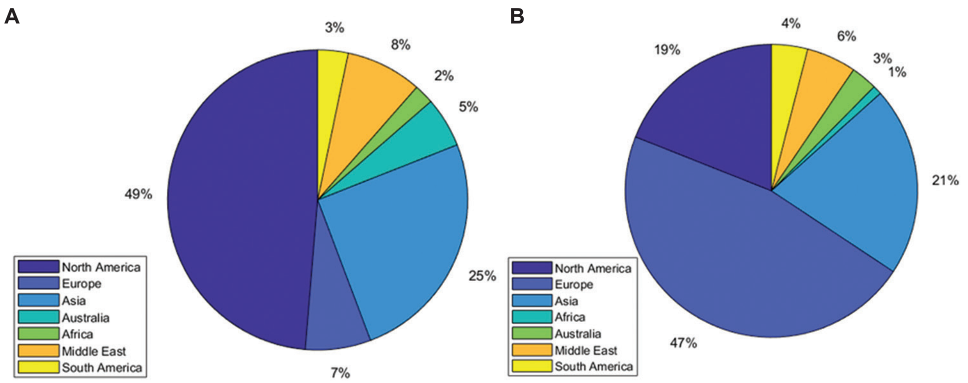

Figure 3. Geographical distribution of papers related to artificial intelligence (AI) in medical (med) education (ed) and XR based on the Web of Science. (A) Distribution of papers on AI+med in radiology. (B) Distribution of papers on AI+med in Extended Reality (i.e., virtual reality, augmented reality, mixed reality, and metaverse).

tool. Similarly, Andersson *et al.*[15] and Mosh *et al.*[16] concluded that AI helps reduce physicians' workload. In addition, during the COVID-19 pandemic, distance learning developed significantly and came to cover the whole world, not just people living in remote and inaccessible areas. This was made possible by the development of technology, including AI-based solutions. Furthermore, the high cost of practical offline classes, especially in the field of medicine, and the consequently limited opportunities for participation make remote learning solutions appear to be an accessible and natural development of the educational sector. Indeed, the potential of AI in distance education has been demonstrated by Garlinska *et al.*,[17] who pointed out that AI-based algorithms can be applied to the personalization of learning content, automated grading, and virtual touring, including as they do features such as speech-to-text and text-to-speech.

More specifically, AI has the potential to improve the diagnostic process, which is crucial for education in this field.[18] Radiology is a significant area of medicine. In July and August 2023, Gordon *et al.*[19] conducted a thorough search of publications found in PubMed/MEDLINE, EMBASE, and MedEdPublish, looking for keywords connected to the use of AI in medical education. The largest preponderance occurred in the field of radiology (11.2%), followed by surgery (8.7%). Then, Pinto dos Santos *et al.*[20] conducted an anonymous survey concerning the attitude of radiology students to AI applications, whereby 83% of participants believed AI-based algorithms could potentially detect pathological changes in radiological images, while 56% felt they were unable to correctly interpret these changes. At the same time, 68% of respondents admitted they were not aware of the technologies and risks associated with AI implementation in radiology. This indicates a significant gap among medical staff regarding AI technology, which may hinder its effective use.

Visualization constitutes a major area where AI-based solutions play a significant role in medical education. These solutions enable the creation of realistic 3D visualizations of organs, their abnormalities, and the entire human body. They can then be used to develop virtual reality (VR) medical simulators in the metaverse and medical holograms (mixed reality).[21-23] Thus, AI has an integral role in the development of the virtual environment. Similarly, surgical simulation provides physicians with richer, more realistic tools for training and improving their skills.[24] It enables the production of 3D interactive objects from medical data such as magnetic resonance imaging, computer tomography, and other techniques.[25,26] This helps students understand medical content, as demonstrated in a pilot study by Sariciliar *et al*.[27] The use of solutions based on AI and XR also allows the operating table to be observed from an unlimited perspective and from anywhere in the world without disrupting the course of the operation (although this is naturally dependent on a stable internet connection.[28,29] Furthermore, the attending physician also has the opportunity to consult with another physician in a different location without leaving the patient and the operating table. Operations can also be recorded and played back from different observation perspectives, which also has a significant educational dimension. An interesting AI-based application field has also been pointed out by Winker-Schwartz *et al*.,[30] namely the use of machine learning to analyze texts to assess knowledge in the area of surgery. AI can also help assess the level of knowledge and skills of the user performing surgery in a virtual environment.[31] For this purpose, electroencephalography recording was used as input data. It turned out that the records differ depending on qualifications and experience, and this can be used to develop a classifier of their skills (markers of experience). However, a major contrast is to be made between the application of AI in medical practice in high and low-income countries, where the latter find the requirement for expensive equipment and high-capacity internet connectivity challenging.

In contrast to the visual side, text is a vast area where AI has the potential for extensive use in many use cases. A GenText engine such as ChatGTP based on large language models (LLMs) is capable of quickly producing large amounts of convincing natural language text of many types and purposes, from chatbots[32] to scientific writing.[33] Although its use as a production tool has led to bans and restrictions in the higher education sector due to fears that students will lose the capability of writing original work or thinking for themselves, the sector is now beginning to accept that there will be a place for ChatGPT and it is better to educate users than to ban it outright.[34] Apart from a tool to assist in the production of text (ranging from reports, notes, and summaries to discursive pieces), ChatGPT can be used to create learning programs themselves. The potential of AI in distance education has been demonstrated by Garlinska *et al*.,[17] who pointed out that AI-based algorithms can be applied to the personalization of learning content, automated grading, and virtual touring, including as they do features such as speech-to-text and text-to-speech. Then, GenText can offer an AI-human interactive experience on top. For example, Beilby *et al*.[35] evaluated the use of ChatGTP for providing fertility information that could inform decision-making. They found that while ChatGPT effectively sorted through a large amount of information and generated summaries, medical practitioners were still needed for further explanations and counseling. Moreover, Funk and coauthors demonstrated statistically significant differences between ChatGPT versions.[36] Specifically, ChatGPT-4 was found to be more consistent, providing 44.9% more correct answers to medical questions than version 3.5.

While AI-based tools such as ChatGTP offer many benefits, they are not without drawbacks.[37] An undoubted advantage is the fact that ChatGPT is a tool that provides the user with output at any time; the only requirement is to have an account and good internet access. Moreover, through the way the user asks questions and based on the context of the conversation, ChatGTP adapts the content to its interlocutor.[38]

There is also an option to include in the question the level at which the user would like to get an answer. This works well even when working on the basic free level rather than the paid (Doctor GPT). However, the biggest limitation is the strict dependence on the nature of the input data applied. In the case of ChatGTP, it is also worth noting the fact that databases have not been updated since January 2022, which is an important issue considering the rapid development of medicine. Furthermore, this operation results from the very principles of its internal structure, its ability to create answers based on patterns within data sequences. In addition, it may provide incomplete information,[39] which may cause ChatGPT to create literature sources that do not exist in reality. Another significant disadvantage of LLMs is the lack of consideration for the context of medical concepts or the nuances of patient care. In addition, there is an ethical issue related to the security of patient data. A further limitation is the fact that ChatGPT operates in English, but data added in other languages could be converted into English by automated translation. In the future, this can improve communication with both patients and students who do not speak English fluently. Furthermore, a limitation is the fact that ChatGPT does not have a built-in data reliability assessment module. It is also possible to overtrain LLMs. It is worth noting that ChatGPT is a content aggregator and analyzer based on a language model, not a source of knowledge in itself. It should rather be treated as a supporting tool or kind of guide whose answers should be examined critically. On the other hand, combining ChatGPT with another tool, such as virtual simulators, can be extremely beneficial for medical students.[40] However, it is during this time that ChatGPT should be thoroughly tested against possible errors that can be made in medical education processes. It is also worth emphasizing that the long-term impact of AI tools, including ChatGPT, on learning outcomes, especially in the field of medicine, should be examined.[41]

On the other hand, an interesting study[42] analyzed medical students' readiness for AI-based solutions. The findings revealed that students who believed AI technologies would contribute to their profession and reduce workload outnumbered those who held a different view. In addition, a study[43] proposed a Persian version of the Medical AI Readiness Scale to evaluate the readiness of medical students to work with AI, including factors, such as cognition, ability, vision, and ethics.

4. Ethical risks in the implementation of AI in medical education

Each of the four examples of AI's significant role in medicine and medical education offers great hope for rapid improvements in medical practice. However, these advancements come with ethical risks that, if not addressed, could result in a curse of malpractice and bad outcomes for educationalists and their students as well as for practitioners. There has been a discussion regarding AI and ethics for many years, as illustrated by Dennett's vision of a novel-writing machine and the dilemmas it raises about the notion of self.[44] Yet, it is only recently that a focus on ethical risks, AI, and medical education has appeared, no doubt in tandem with the rapid development of technology. Indeed, on the general level, as noted above, Weidener and Fischer[13] demonstrated that there is a lack of discussion concerning AI and medical education overall, even though, as Civaner *et al.*[14] pointed out, there is a recognition amongst many medical students that AI needs to play a role in medical education. This shows that there is a student (or consumer) demand for AI in educational curricula and a need for educators to fill that gap. There is thus a clear requirement for AI to be integrated into medical education programs, but reasons can be advanced for the slow pace of adoption. For example, such programs are extensive and well-established, and there may be resistance from course designers and managers, educators, and other stakeholders.[45] On the other hand, the integration of AI into medical education is likely inevitable, paving the way for serious disruption and commercial opportunities. Indeed, it is necessary since a lack of integration will constitute a further type of broad ethical risk: if students are not equipped with AI knowledge, they will be less able to cope with the various and detailed types of ethical risk as practitioners. However, advances are being made even while calls for a faster pace of change are being made.[46,47] An outline model for the application of AI in medical education is provided by Zarei *et al.*,[48] along with an assessment of challenges such as the current lack of infrastructure. Krive *et al.*[49] designed and tested a model comprising a modular 4-week AI course, which proved to be successful.

As a specific area, radiology, for example, depends heavily on data.[50-52] It is immediately apparent that the successful manipulation of information-intensive radiological data using AI requires significant computational resources. This raises concerns about energy use, costs, and environmental impact, where developing countries may be at a disadvantage, thus increasing ethical risk for them. Another extremely important issue concerns how the accuracy of AI predictions using various types of metrics is to be evaluated.[53] This is connected with algorithmic fairness:[54] If one method of evaluation produces a different metric than another, the outcome could result in being unfair to one or another cohort, an ethical issue. The most popular algorithms in the field of medicine are the Dice coefficient and accuracy.[4,55] However, there is no accepted standardization for the assessment of such algorithms in medicine. Turning to the issue of data biases, the extensive account provided by Ueda *et al.*[56] broadly separated into machine and human-originated, and the discussion of biases identified by Pregowska and Perkins[5] (passim) prompts the need for two underlying dimensions of bias to be highlighted in addition. The first is intentional and unintentional. The introduction of bias into a dataset (such as the over-representation of one demographic cohort at the expense of another or incorrect,[57] and model and interpretation bias[58]) may be intentional on the part of the human agent or unintentional (due to accident, neglect, human error, or subconscious attitude). Once intentional bias has been identified, the question of motivation arises as a second underlying dimension. Bias can be introduced into dataset selection, and datasets can be manipulated due to social and political attitudes in some societies. The profit motive may also raise issues of control, ownership, deployment, and use of data, and even falsification.[59] The increasing role of AI, along with its ability to create and amplify biases or distort information – complicated by the need for radiological data between institutions and across borders[60] – highlights the importance of transparently identifying agents within the system and their access to AI tools. This transparency should be integrated into medical education from the outset.[61] In addition, convincing practitioners of the significant benefits AI offers to radiology[62] presents serious challenges to those engaged in curriculum and syllabus design. Moreover, there is an external dimension of malign intention represented by cybersecurity threats. Medicine is under increasing attack, and practically no field is more greatly exposed than radiology.[63] Cyber-attacks can range from malicious insider activity to data theft, credential harvesting, and phishing. They can occur at various points in the radiological landscape, including medical devices, wireless systems, data warehouses, and social networks, and the increasing use of AI on both sides has created vulnerabilities.[63] Moreover, there is also no clear overview of approved AI-based medical devices. This leads to inconsistency and increased ethical risk. However, the problem is recognized, and investigations are currently underway by the Food and Drug Administration in the USA,[64,65] and the Medicines and Healthcare Products Regulatory Agency,[65] which is developing guidelines for such devices.[66] Here, broadly understood, cybersecurity is an important issue.[5] AI systems are vulnerable to adversarial attacks,[67] such as the introduction of minor modifications to input data in changing training labels that lead to invalid predictions. Each such attack is a breach of sensitive patient information, and any wrong decision in the medical field has potentially disastrous consequences. This vulnerability extends not only to patient data but also to student data. Tsai and Lin[68] proposed a procedure to evaluate the resistance of AI models based on medical images against these attacks. There are various techniques to defend against adversarial attacks, including data augmentation, adversarial training, and robust optimization. However, establishing effective protection protocols remains a challenge.[69]

In the development of two further contrasting and specific areas of AI, there is evidence of AI being used in education, VR, and GenText.[70] In the case of VR, although some evidence of adoption has been found to be sparse, as in the database search by Lie *et al.*[71] covering November and December 2021, a subsequent more extensive literature search study in the period January 2017 to March 2022 demonstrated a rapid and increasing take-up perhaps in the latter part of this period, although this is not stated in the research.[72] Moreover, students trained using VR produce better results than those conventionally taught. Kim and Kim[73] identified and examined 24 studies and a sub-group of 18 on the use versus non-use of VR in medical education and found that "there was a significant improvement in the VR group's skill and satisfaction levels, and that less immersive VR was more efficacious for knowledge outcomes than fully immersive VR" (ibid., p. 13). Greater student satisfaction in using AI is also confirmed by Leng,[74] who found that in the case of learning anatomy, ChatGPT has increased student engagement. Then, in a small-scale study of 44 students aimed at validating VR-based medical training, Pedram *et al.*[75] not only found a user acceptance level of 75% but also an outperformance by those using VR of the control group that did not. These studies reinforce the view that there is a greater ethical risk in a sluggish implementation of AI in medical education than in a rapid one. Slow implementation will result in inferior education. In turn, this will lead to slower and possibly deficient deployment of AI in the clinic, with consequently worse patient outcomes. While the fast deployment of AI in medical education will bring lower ethical risk, another aspect of risk may be avoided, that of the vulnerability of data. In a clinical setting, real patient data will be used. In a VR scenario, simulated data are sufficient. Mergen *et al.*[76,77] have developed a project tool entitled "medical tr.AI.ning," an immersive VR learning platform based on AI that generates simulated patient data, thus obviating ethical concerns.

Regarding ChatGPT and other potential GenText engines, there are many points of ethical risk in medical education. Once more, the output quality depends on the input datasets. Very often, data, especially medical data, is burdened with various types of bias.[46] There is also a further question of whether ChatGPT is biased as a collection of algorithms or whether algorithmic bias could be introduced unethically.[78] If bias can occur at these two levels, there is a further systemic ethical threat in the vulnerability of GenText and other engines to jailbreak, where an AI system acts outside the restrictions placed on it by its designers.[79] Further alarming consequences may arise when an AI that has broken free can create other AIs that may produce harmful output, such as producing a set of instructions for synthesizing methamphetamine.[80] However, it is at the day-to-day level that ChatGPT causes a great deal of concern: at face value, ChatGPT can be used to generate substantial amounts of convincing text. Such text can be used for framework and content infill for curricula and syllabi (by course managers and designers), teaching material (by educators), and assignments (by students). However, where there is a risk that the content generated is at risk of being out of date (depending at least on the latency of input protocols and difficulty in ensuring the provision of the latest academic material (due to secrecy concerns) and that ChatGPT is capable of hallucinating with it comes to references,[81] the validity of the output will be variable and at times questionable. In addition, Májovský *et al.*, 2023[82] considered ChatGTP as a tool for the generation of fake medical papers. The whole process took an hour, and it turned out that the text looked convincing. Although references and specific errors raised doubts, these errors could only be detected by an experienced reader, here a medical doctor. This creates a risk that a student using ChatGTP or a student generating, for example, text for his or her work will not be able to find irregularities that may be important. Chio *et al*.[83] put forward the same argument, raising ChatGPT's lack of critical reflection in the case of nurse education. ChatGTP does not act in such a way as to assess the credibility of sources; when asked to provide literature, in many cases, it creates references that do not exist. This all amounts to a substantial ethical risk. Simply put, the output from GenText cannot be completely relied upon and needs human moderation. In a study conducted by Alam *et al*.[84] on writing in an undergraduate medical degree course in Darussalam, they found substantial inclusion of ChatGPT-generated text and citations to non-existent references. As a result, they propose that educators should be more aware of Gen Text detection tools. Another ethical issue arising here concerns the fact that ChatGPT can give answers to students very quickly: The normal study processes of critical thinking (assessing information, making connections, and drawing conclusions) are thus bypassed. Another ethical issue is the possibility of students writing potentially fraudulent assignments. This indicates the need for teaching how ChatGPT can be used appropriately in the learning situation. Apart from discursive documents and reports, there are many other types of text that can be generated. These include personalized learning plans and treatment plans. If these are relied on without scrutiny or moderation, adverse results could occur, such as misdiagnosis and bad treatment (of oneself and others). This needs to be regulated on a high level and accounted for in local ethical policies and educational practices,[85] and calls are emerging for the development of new educational governance of AI in higher medical education.[86]

The existing system of text types and uses provides a benchmark against which to assess ChatGPT, but the components of that benchmark are not 100% accurate, acceptable, correct, or free from contestable interpretation. Indeed, the production of error (however defined) and the principle of falsifiability are a necessary condition for the advancement of knowledge.[87] In that case, if ChatGPT is to be criticized as falling short, as it does, to what extent can it serve as a useful tool? The immediate answer lies in a case-by-case detailed evaluation and benchmarking process where each instance is allocated a point on a scale of usefulness and risk. Specific cases include that discussed by Abdelhady and Davis,[88] who reported that ChatGPT was able to record operative notes extremely quickly and to a high level of accuracy compared with manual procedures and was deemed acceptable by surgeons and patients alike. Furthermore, several research studies have been conducted where ChatGPT was required to take a variety of medical tests, such as the UK BMAT, TMUA, LNAT, and TSA examinations, the United States Medical Licensing Exam (USMLE), and certain university tests.[5] In all cases, ChatGPT came out sufficiently well to be deemed able to set and mark tests, although Giannos and Delardas[89] found that it had a poor knowledge of science and mathematics. Indeed, in many areas of medical education, this tool does not offer specialized knowledge, as in the case of pediatric cardiology education.[90] In a further study, Danesh *et al*. (2024) tested both the free and the premium versions of ChatGTP in terms of its ability to pass professional examinations (excluding questions containing imaging data). ChatGTP was able to answer 50% of the questions correctly in the free version and 70% of the questions in the paid version. Similar examination results were obtained in the fields of orthopedics[91] and health professional exams.[92] Finally, Sevgi *et al*.[93] proposed an evaluation of ChatGTP in the field of neurosurgery by asking it to create questions at the level of a neurosurgery board exam. The question format was to be multiple choices and the answers were also to be generated. Next, it was asked to devise artificial neurosurgical cases with examinations and treatment histories. The final stage involved an evaluation of the tool's ability to create articles in this area. It turned out that the proposed cases did serve to help neurosurgery students develop their knowledge. However, it transpired that a correct assessment of the solutions proposed by ChatGTP was only possible under the supervision of a person with appropriate medical knowledge, in this case, an experienced neurosurgeon.

5. Conclusion

The application of AI allows the efficient analysis of huge amounts of data in a finite time. It can be considered a powerful computational tool for solving complex problems related to pattern recognition, classification, grouping, behavior prediction, or, more generally, approximation of functions and processes. Consequently, AI is becoming a highly precise tool in medicine. It is worth stressing that compared to statistical methods; it is more susceptible to various types of threats in comparison, which is a result of its complexity, data dependence, and susceptibility to adversarial attacks. Although AI offers many benefits in medical education, ethical concerns about its accessibility, validity, use, and implementation raise many questions. AI can be implemented in medical education in a variety of beneficial and relatively uncontroversial ways. These include the rapid analysis of large-scale simulated datasets (thus obviating requirements of real-life patient data regulation), pattern recognition and diagnostics (as in radiology), general educational support in the design of personalized learning programs (at least on a basic level), models built in VR as teaching aids, and uses of GenText such as rapid assembly of post-operative notes.

On the other hand, barriers to implementation and use in developing countries, such as limited internet connectivity, have resulted in lower levels of discussion around global fairness as an ethical issue. In addition, input data latency and potential dataset and algorithmic bias raise genuine concerns about the output validity, especially regarding GenText and statistical analytical output. A particular concern in discursive production is the ability of GenText to hallucinate (create non-existent references) and create output text of a biased nature (opinions and accounts ultimately derived from bias found in input datasets and algorithmic structures) that could distort the nature of medical education, leading to bad ethical and practical outcomes in the future. Furthermore, the intensive development of computer hardware, including quantum computers, and the algorithms themselves, and in particular their learning methods, which is the heart of AI, is likely to significantly shorten the time needed for more precise analysis, which is crucial in the context of medical data.

Acknowledgments

None.

Funding

This work was partially supported by the National Center for Research and Development (research grant Infostrateg I/0042/2021-00).

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: All authors
Writing – original draft: All authors
Writing – review & editing: All authors

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Not applicable.

Further disclosure

None.

References

1. Dolega-Dolegowski D, Proniewska K, Dolega-Dolegowska M, *et al*. Application of holography and augmented reality based technology to visualize the internal structure of the dental root – a proof of concept. *Head Face Med*. 2022;18(1):12.

   doi: 10.1186/s13005-022-00307-4

2. Baker RS, Hawn A. Algorithmic bias in education. *Int J Artif Intell Educ*. 2022;32:1052-1092.

   doi: 10.1007/s40593-021-00285-9

3. Albahri AS, Duhaim AM, Fadhel MA, *et al*. A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Inform Fusion*. 2023;96:156-191.

   doi: 10.1016/J.INFFUS.2023.03.008

4. Rudnicka Z, Szczepanski J, Pregowska A. Artificial intelligence-based algorithms in medical image scan segmentation and intelligent visual content generation-a concise overview. *Electronics* (*Basel*). 2024;13(4):746.

   doi: 10.3390/electronics13040746

5. Pregowska A, Perkins M. *Artificial Intelligence in Medical Education: Technology and Ethical Risk*. Available from: https://ssrn.com/abstract=4643763 [Last accessed on 2024 Oct 18].

   doi: 10.2139/ssrn.4643763

6. Naik N, Hameed BMZ, Shetty DK, *et al*. Legal and ethical consideration in artificial intelligence in healthcare: Who takes responsibility? *Front Surg*. 2022;9:862322.

   doi: 10.3389/fsurg.2022.862322

7. Bae CY, Im Y, Lee J, *et al*. Comparison of biological age prediction models using clinical biomarkers commonly measured in clinical practice settings: AI techniques Vs. traditional statistical methods. *Front Anal Sci*. 2021;1:709589.

   doi: 10.3389/frans.2021.709589

8. Hassija V, Chamola V, Mahapatra A, *et al*. Interpretability of black-box models: A review on explainable artificial intelligence (XAI). *Cognit Comput*. 2024;16:45-74.

   doi: 10.1007/s12559-023-10179-8

9. Kunze KN, Williams RJ 3rd, Ranawat AS, *et al*. Artificial intelligence (AI) and large data registries: Understanding the advantages and limitations of contemporary data sets for use in AI research. *Knee Surg Sports Traumatol Arthrosc*. 2024;32(1):13-18.

   doi: 10.1002/ksa.12018

10. Baniecki H, Biecek P. Adversarial attacks and defenses in explainable artificial intelligence: A survey. Information Fusion. 2023;107:102303.

    doi: 10.1016/j.inffus.2024.102303

11. Rethlefsen ML, Kirtley S, Waffenschmidt S, *et al*. PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev*.

2021;10:39.

doi: 10.1186/s13643-020-01542-z

12. Lee J, Wu AS, Li D, Kulasegaram K (Mahan). Artificial intelligence in undergraduate medical education: A scoping review. *Acad Med*. 2021;96(11S):S62-S70.

    doi: 10.1097/ACM.0000000000004291

13. Weidener L, Fischer M. Teaching AI ethics in medical education: A scoping review of current literature and practices. *Perspect Med Educ*. 2023;12:399-410.

    doi: 10.5334/pme.954

14. Civaner MM, Uncu Y, Bulut F, Chalil EG, Tatli A. Artificial intelligence in medical education: a cross-sectional needs assessment. *BMC Med Educ*. 2022;22(1):772.

    doi: 10.1186/s12909-022-03852-3

15. Andersson J, Nyholm T, Ceberg C, *et al*. Artificial intelligence and the medical physics profession - a Swedish perspective. *Phys Med*. 2021;88:218-225.

    doi: 10.1016/J.EJMP.2021.07.009

16. Mosch L, Fürstenau D, Brandt J, *et al*. The medical profession transformed by artificial intelligence: Qualitative study. *Digit Health*. 2022;8:20552076221143903.

    doi: 10.1177/20552076221143903

17. Garlinska M, Osial M, Proniewska K, Pregowska A. The influence of emerging technologies on distance education. *Electronics* (*Switzerland*). 2023;12(7):1550.

    doi: 10.3390/electronics12071550

18. Klauschen F, Dippel J, Keyl P, *et al*. Toward explainable artificial intelligence for precision pathology. *Annu Rev Pathol Mech Dis*. 2024;19(1):541-570.

    doi: 10.1146/annurev-pathmechdis-051222-113147

19. Gordon M, Daniel M, Ajiboye A, *et al*. A scoping review of artificial intelligence in medical education: BEME Guide No. 84. *Med Teach*. 2024;46:446-470.

    doi: 10.1080/0142159X.2024.2314198

20. Pinto Dos Santos D, Giese D, Brodehl S, *et al*. Medical students' attitude towards artificial intelligence: A multicentre survey. *Eur Radiol*. 2019;29:1640-1646.

    doi: 10.1007/s00330-018-5601-1

21. Chengoden R, Victor N, Huynh-The T, *et al*. Metaverse for healthcare: A survey on potential applications, challenges and future directions. *IEEE Access*. 2023;PP:1-1.

    doi: 10.1109/ACCESS.2023.3241628

22. Musamih A, Yaqoob I, Salah K, *et al*. Metaverse in healthcare: Applications, challenges, and future directions. *IEEE Consum Electron Mag*. 2023;12(4):33-46.

    doi: 10.1109/MCE.2022.3223522

23. Zechner O, Guirao DG, Schrom-Feiertag H, *et al*. Multimodal technologies and interaction NextGen training for medical first responders: Advancing mass-casualty incident preparedness through mixed reality technology. *Multimodal Technol Interact*. 2023;7:113.

    doi: 10.3390/mti7120113

24. Balak N, Ganau M, Tsianaka E, Park JJ, Tiefenbach J, Demetriades AK. The role of artificial intelligence in surgical simulation. *Front Med Technol*. 2022;4:1076755.

    doi: 10.3389/fmedt.2022.1076755

25. Neves CA, Tran ED, Kessler IM, Blevins NH. Fully automated preoperative segmentation of temporal bone structures from clinical CT scans. *Sci Rep*. 2021;11:116.

    doi: 10.1038/s41598-020-80619-0

26. Hamabe A, Ishii M, Kamoda R, *et al*. Artificial intelligence-based technology to make a three-dimensional pelvic model for preoperative simulation of rectal cancer surgery using MRI. *Ann Gastroenterol Surg*. 2022;6:788-794.

    doi: 10.1002/ags3.12574

27. Saricilar EC, Burgess A, Freeman A, *et al*. A pilot study of the use of artificial intelligence with high-fidelity simulations in assessing endovascular procedural competence independent of a human examiner. *ANZ J Surg*. 2023;93:1525-1531.

    doi: 10.1111/ans.18484

28. Mirchi IN, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS One*. 2020;15:e0229596.

    doi: 10.1371/journal.pone.0229596

29. Proniewska K, Dolega-Dolegowski D, Kolecki R, Osial M, Pregowska A. The 3D operating room with unlimited perspective change and remote support. In: *Applications of Augmented Reality - Current State of the Art [Working Title]*. London: Intechopen; 2023.

    doi: 10.5772/intechopen.1002252

30. Winkler-Schwartz A, Bissonnette V, Mirchi N, *et al*. Artificial intelligence in medical education: Best practices using machine learning to assess surgical expertise in virtual reality simulation. *J Surg Educ*. 2019;76(6):1681-1690.

    doi: 10.1016/J.JSURG.2019.05.015

31. Natheir S, Christie S, Yilmaz R, *et al*. Utilizing artificial intelligence and electroencephalography to assess expertise on a simulated neurosurgical task. *Comput Biol Med*. 2023;152:106286.

    doi: 10.1016/j.compbiomed.2022.106286

32. Wölfel M, Taecharungroj V. "What Can ChatGPT Do?" Analyzing early reactions to the innovative AI Chatbot on twitter. *Big Data Cogn Comput*. 2023;7:35.

doi: 10.3390/bdcc7010035

33. Alkaissi H, Mcfarlane SI. Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*. 2023;15:e35179.

    doi: 10.7759/cureus.35179

34. Generative AI Tools for USD Law Students. University of San Diego Legal Research Center Guides. Available from: https://lawlibguides.sandiego.edu/c.php?g=1317323&p=9686671 [Last accessed on 2024 Oct 18].

35. Beilby K, Hammarberg K. ChatGPT: A reliable fertility decision-making tool? *Hum Reprod*. 2024;39:443-447.

    doi: 10.1093/humrep/dead272

36. Funk PF, Hoch CC, Knoedler S, *et al*. ChatGPT's response consistency: A study on repeated queries of medical examination questions. *J Investig Health Psychol Educ*. 2024;14(3):657-668.

    doi: 10.3390/ejihpe14030043

37. Mu Y, He D. The potential applications and challenges of ChatGPT in the medical field. *Int J Gen Med*. 2024;17:817-826.

    doi: 10.2147/ijgm.s456659

38. Carr SE, Canny BJ, Wearn A, *et al*. Twelve tips for medical students experiencing an interruption in their academic progress. *Med Teach*. 2022;44(10):1081-1086.

    doi: 10.1080/0142159X.2021.1921134

39. Shen Y, Heacock L, Elias J, *et al*. ChatGPT and other large language models are double-edged swords. *Radiology*. 2023;307(2):e230163.

    doi: 10.1148/radiol.230163

40. Lee H. The rise of ChatGPT: Exploring its potential in medical education. 2023.

    doi: 10.1002/ase.2270

41. 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

42. Emir B, Yurdem T, Ozel T, *et al*. Artificial intelligence readiness status of medical faculty students. *Konuralp Med J*. 2024;16(1):88-95.

    doi: 10.18521/ktd.1387826

43. Rezazadeh H, Ahmadipour H, Salajegheh M. Psychometric evaluation of Persian version of medical artificial intelligence readiness scale for medical students. *BMC Med Educ*. 2023;23(1):527.

    doi: 10.1186/s12909-023-04516-6

44. Dennett D. *The Self as a Center of Narrative Gravity. Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ: Lawrence Erlbaum; 1992. p. 53.

45. Acharya V, Padhan P, Bahinipati J, et al. Artificial intelligence in medical education. *J Integr Med Res*. 2023;1(3):87.

    doi: 10.4103/jimr.jimr_17_23

46. Katznelson G, Gerke S. The need for health AI ethics in medical school education. *Adv Health Sci Educ*. 2021;26:1447-1458.

    doi: 10.1007/s10459-021-10040-3

47. Ötleş E, James CA, Lomis KD, Woolliscroft JO. Teaching artificial intelligence as a fundamental toolset of medicine. *Cell Rep Med*. 2022;3(12):100824.

    doi: 10.1016/J.XCRM.2022.100824

48. Zarei M, Eftekhari Mamaghani H, Abbasi A, Hosseini MS. Application of artificial intelligence in medical education: A review of benefits, challenges, and solutions. *Med Clín Práct*. 2024;7(2):100422.

    doi: 10.1016/J.MCPSP.2023.100422

49. Krive J, Isola M, Chang L, Patel T, Anderson M, Sreedhar R. Grounded in reality: Artificial intelligence in medical education. *JAMIA Open*. 2023;6:ooad037.

    doi: 10.1093/jamiaopen/ooad037

50. Piorkowski A, Obuchowicz R, Najjar R. Redefining radiology: A review of artificial intelligence integration in medical imaging. *Diagnostics* (*Basel*). 2023;13:2760.

    doi: 10.3390/diagnostics13172760

51. Brady AP, Allen B, Chong J, Kotter E, Kottler N, Mongan J, *et al*. Developing, purchasing, implementing and monitoring AI tools in radiology: practical considerations. A multi-society statement from the ACR, CAR, ESR, RANZCR & RSNA. Insights Imaging. 2024;15(1):16.

    doi: 10.1186/s13244-023-01541-3

52. Choudhury A, Elkefi S. Acceptance, initial trust formation, and human biases in artificial intelligence: Focus on clinicians. *Front Digit Health*. 2022;4:966174.

    doi: 10.3389/fdgth.2022.966174

53. Pagano TP, Loureiro RB, Lisboa FVN, *et al*. Bias and unfairness in machine learning models: A systematic review on datasets, tools, fairness metrics, and identification and mitigation methods. *Big Data Cogn. Comput*. 2023;7:15.

    doi: 10.3390/bdcc7010015

54. Pessach D, Shmueli E. A review on fairness in machine learning. *ACM Comput Surv*. 2022;55(3):51.

    doi: 10.1145/3494672

55. Rudnicka Z, Proniewska K, Perkins M, Pregowska A. Cardiac healthcare digital twins supported by artificial intelligence-based algorithms and extended reality-a systematic review. *Electronics* (*Basel*). 2024;13(5):866.

    doi: 10.3390/electronics13050866

56. Ueda D, Kakinuma T, Fujita S, *et al*. Fairness of artificial intelligence in healthcare: Review and recommendations. *Jpn J Radiol*. 2024;42:3-15.

doi: 10.1007/s11604-023-01474-3

57. Shen D, Liu T. Grand challenges in AI in radiology. *Front Radiol.* 2021;1:629992.

doi: 10.3389/fradi.2021.629992

58. Park Y, Hu J. Bias in artificial intelligence: Basic primer. *Clin J Am Soc Nephrol*. 2023;18(3):394-396.

doi: 10.2215/CJN.0000000000000078

59. Khosravi P, Schweitzer M. Artificial intelligence in neuroradiology: A scoping review of some ethical challenges. Front Radiol. 2023;3:1149461.

doi: 10.3389/fradi.2023.1149461

60. Bell LC, Shimron E. Sharing data is essential for the future of AI in medical imaging. *Radiol Artif Intell*. 2023;6(1):e230337.

doi: 10.1148/ryai.230337

61. Bernstein MH, Atalay MK, Dibble EH, *et al*. Can incorrect artificial intelligence (AI) results impact radiologists, and if so, what can we do about it? A multi-reader pilot study of lung cancer detection with chest radiography. *Eur Radiol*. 2023;33(11):8263-8269.

doi: 10.1007/s00330-023-09747-1

62. Eltawil FA, Atalla M, Boulos E, Amirabadi A, Tyrrell PN. Analyzing barriers and enablers for the acceptance of artificial intelligence innovations into radiology practice: A scoping review. *Tomography*. 2023;9(4):1443-1455.

doi: 10.3390/tomography9040115

63. Kelly BS, Quinn C, Belton N, *et al*. Cybersecurity considerations for radiology departments involved with artificial intelligence. *Eur Radiol*. 2023;33:8833-8841.

doi: 10.1007/s00330-023-09860-1

64. U.S. Food & Drug Administration. Artificial intelligence-enabled medical devices. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices [Last accessed on 2024 Oct 08].

65. Medicines & Healthcare Products Regulatory Agency. Available from: https://www.gov.uk/government/organisations/medicines-and-healthcare-products-regulatory-agency[Last accessed on 2024 Oct 08].

66. UK Digital Health - the Future of Software as a Medical Device. Available from: https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap [Last accessed on 2024 Oct 18].

67. Khazane H, Ridouani M, Salahdine F, Kaabouch N. A holistic review of machine learning adversarial attacks in IoT networks. *Future Internet*. 2024;16:32.

doi: 10.3390/fi16010032

68. Tsai MJ, Lin PY. Medical images under tampering. *Multimed Tools Appl*. 2024;83(24):65407-65439.

doi: 10.1007/s11042-023-17968-1

69. Prajapati JB, Kumar A, Singh S, *et al*. Artificial intelligence-assisted generative pretrained transformers for applications of ChatGPT in higher education among graduates. *SN Soc Sci*. 2024;4:19.

doi: 10.1007/s43545-023-00818-0

70. Narayanan S, Ramakrishnan R, Durairaj E, Das A. Artificial intelligence revolutionizing the field of medical education. *Cureus*. 2023;15(11):e49604.

doi: 10.7759/cureus.49604

71. Lie SS, Helle N, Sletteland NV, Dubland Vikman M, Bonsaksen T. Implementation of virtual reality in health professions education: Scoping Review. *JMIR Res Protoc*. 2022;11:e37222.

doi: 10.2196/37222

72. Dhar E, Upadhyay U, Huang Y, *et al*. A scoping review to assess the effects of virtual reality in medical education and clinical care. *Digit Health*. 2023;9:20552076231158022.

doi: 10.1177/20552076231158022

73. Kim HY, Kim EY. Effects of medical education program using virtual reality: A systematic review and meta-analysis. *Int J Environ Res Public Health*. 2023;20(5):3895.

doi: 10.3390/ijerph20053895

74. Leng L. Challenge, integration, and change: ChatGPT and future anatomical education. *Med Educ Online*. 2024;29(1):2304973.

doi: 10.1080/10872981.2024.2304973

75. Pedram S, Kennedy G, Sanzone S. Assessing the validity of VR as a training tool for medical students. *Virtual Real*. 2024;28:15.

doi: 10.1007/s10055-023-00912-x

76. Mergen M, Meyerheim M, Graf N. Reviewing the current state of virtual reality integration in medical education - a scoping review protocol. *Syst Rev*. 2023;12(1):97.

doi: 10.1186/s13643-023-02266-6

77. Mergen M, Meyerheim M, Graf N. Towards integrating virtual reality into medical curricula: A single center student survey. *Educ Sci*. 2023;13:477.

doi: 10.3390/educsci13050477

78. Heikkilä M. AI language models are rife with different political biases. *MIT Tech Rev*. Published August 7, 2023. Available from: https://www.technologyreview.com/2023/08/07/1077324/ai-language-models-are-rife-

with-political-biases [Last accessed on 2024 Oct 08].

79. Feng Y, Chen Z, Kang Z, *et al*. JailbreakLens: Visual analysis of jailbreak attacks against large language models. *arXiv*. Preprint posted online 2024.

doi: 10.48550/ARXIV.2404.08793

80. Shah R, Feuillade-Montixi Q, Pour S, Tagade A, Casper S, Rando J. Scalable and transferable black-box jailbreaks for language models via persona modulation. *arXiv*. Preprint posted online 2023.

doi: 10.48550/arXiv.2311.03348

81. Eysenbach G. The role of ChatGPT, generative language models, and artificial intelligence in medical education: A conversation with ChatGPT and a call for papers. *JMIR Med Educ*. 2023;9:e46885.

doi: 10.2196/46885

82. Májovský M, Černý M, Kasal M, Komarc M, Netuka D. Artificial intelligence can generate fraudulent but authentic-looking scientific medical articles: Pandora's box has been opened. *J Med Internet Res*. 2023;25:e46924.

doi: 10.2196/46924

83. Choi EPH, Lee JJ, Ho MH, Kwok JYY, Lok KYW. Chatting or cheating? The impacts of ChatGPT and other artificial intelligence language models on nurse education. *Nurse Educ Today*. 2023;125:105796.

doi: 10.1016/J.NEDT.2023.105796

84. Alam F, Lim MA, Zulkipli IN. Integrating AI in medical education: embracing ethical usage and critical understanding. *Front Med*. 2023;10:1279707.

doi: 10.3389/fmed.2023.1279707

85. Kitamura FC. ChatGPT is shaping the future of medical writing but still requires Human Judgment. *Radiology*. 2023;307(2):e230171.

doi: 10.1148/radiol.230171

86. Filgueiras F. Artificial intelligence and education governance. *Educ Citizsh Soc Justice*. 2023.

doi: 10.1177/17461979231160674

87. Taran S, Adhikari NKJ, Fan E. Falsifiability in medicine: What clinicians can learn from Karl Popper. *Intensive Care Med*. 2021;47:1054-1056.

doi: 10.1007/s00134-021-06432-z

88. Abdelhady AM, Davis CR. Plastic surgery and artificial intelligence: How ChatGPT improved operation note accuracy, time, and education. *Mayo Clin Proc Digit Health*. 2023;1(3):299-308.

doi: 10.1016/J.MCPDIG.2023.06.002

89. Giannos P, Delardas O. Performance of ChatGPT on UK standardized admission tests: Insights from the BMAT, TMUA, LNAT, and TSA examinations. *JMIR Med Educ*. 2023;9:e47737.

doi: 10.2196/47737

90. Gritti MN, Hussain A, Farid P, Morgan CT. Progression of an artificial intelligence chatbot (ChatGPT) for pediatric cardiology educational knowledge assessment. *Pediatr Cardiol*. 2024;45:309-313.

doi: 10.1007/s00246-023-03385-6

91. Rizzo MG, Cai N, Constantinescu D. The performance of ChatGPT on orthopaedic in-service training exams: A comparative study of the GPT-3.5 turbo and GPT-4 models in orthopaedic education. *J Orthop*. 2024;50:70-75.

doi: 10.1016/J.JOR.2023.11.056

92. Davies NP, Wilson R, Winder MS, *et al*. ChatGPT sits the DFPH exam: Large language model performance and potential to support public health learning. *BMC Med Educ*. 2024;24(1):57.

doi: 10.1186/s12909-024-05042-9

93. Sevgi UT, Erol G, Doğruel Y, *et al*. The role of an open artificial intelligence platform in modern neurosurgical education: A preliminary study. *Neurosurg Rev*. 1998;46:86.

doi: 10.1007/s10143-023-01998-2