

Viewpoint

Ethical Considerations and Fundamental Principles of Large Language Models in Medical Education: Viewpoint

Li Zhui, PhD; Li Fenghe, PhD; Wang Xuehu, PhD; Fu Qining, PhD; Ren Wei, PhD

Department of Vascular Surgery, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

Corresponding Author:

Ren Wei, PhD

Department of Vascular Surgery

The First Affiliated Hospital of Chongqing Medical University

No. 1 of Youyi Road, Yuzhong District

Chongqing, 400016

China

Phone: 86 13658339771

Fax: 86 89011231

Email: renwei_2301@yeah.net

Abstract

This viewpoint article first explores the ethical challenges associated with the future application of large language models (LLMs) in the context of medical education. These challenges include not only ethical concerns related to the development of LLMs, such as artificial intelligence (AI) hallucinations, information bias, privacy and data risks, and deficiencies in terms of transparency and interpretability but also issues concerning the application of LLMs, including deficiencies in emotional intelligence, educational inequities, problems with academic integrity, and questions of responsibility and copyright ownership. This paper then analyzes existing AI-related legal and ethical frameworks and highlights their limitations with regard to the application of LLMs in the context of medical education. To ensure that LLMs are integrated in a responsible and safe manner, the authors recommend the development of a unified ethical framework that is specifically tailored for LLMs in this field. This framework should be based on 8 fundamental principles: quality control and supervision mechanisms; privacy and data protection; transparency and interpretability; fairness and equal treatment; academic integrity and moral norms; accountability and traceability; protection and respect for intellectual property; and the promotion of educational research and innovation. The authors further discuss specific measures that can be taken to implement these principles, thereby laying a solid foundation for the development of a comprehensive and actionable ethical framework. Such a unified ethical framework based on these 8 fundamental principles can provide clear guidance and support for the application of LLMs in the context of medical education. This approach can help establish a balance between technological advancement and ethical safeguards, thereby ensuring that medical education can progress without compromising the principles of fairness, justice, or patient safety and establishing a more equitable, safer, and more efficient environment for medical education.

(*J Med Internet Res* 2024;26:e60083) doi: [10.2196/60083](https://doi.org/10.2196/60083)

KEYWORDS

medical education; artificial intelligence; large language models; medical ethics; AI; LLMs; ethics; academic integrity; privacy and data risks; data security; data protection; intellectual property rights; educational research

Introduction

Large language models (LLMs), such as OpenAI's ChatGPT series, Microsoft's Copilot, and Google's Gemini, are sophisticated artificial intelligence (AI) tools that are trained on vast text datasets via deep learning techniques. These models exhibit the ability to generate language that is highly similar to human speech; thus, their use in medicine is becoming increasingly widespread. From assisting scientific research to informing clinical decisions and supporting medical education,

LLMs can be used to optimize the use of information and resources across diverse domains [1].

In the technology-driven transformation of medical education, the introduction of LLMs has the potential to revolutionize traditional learning and teaching methods. These advanced language models can significantly enhance the personalization and practicality of education by generating tailored learning materials and simulating digital experiences that resemble real clinical scenarios [2]. LLMs dynamically generate customized

learning materials and assessments based on students' progress and needs, thereby improving the efficiency and effectiveness of learning [3]. Moreover, by simulating patient interactions, LLMs can help students practice and hone their diagnostic and communication skills in a safe environment, which is a crucial element of attempts to nurture the practical abilities of medical professionals [4].

Biomedical ethics is founded on 4 fundamental principles: autonomy, nonmaleficence, beneficence, and justice [5]. Medical education should not only aim to cultivate these principles among students but also continuously evolve and improve under their guidance [6]. Due to the increasingly widespread use of LLMs in medical education research, some foreseeable ethical issues present challenges to these basic principles. For example, the privacy and data security risks associated with LLMs can infringe upon individuals' autonomy to control their own information, thus potentially causing harm. Biases in training data and algorithms can lead to unfair diagnostic and treatment decisions that can harm patients, and these biases may also risk spreading misinformation, thereby violating the principle of beneficence. Additionally, issues related to academic integrity and concerns regarding responsibility and copyright may risk violating the principle of justice, among others.

While previous research on education-related LLMs has emphasized the tasks of enhancing key competencies or exploring educational LLM programs and related concepts, it has failed to account for ethical principles specifically in the development and application of such technologies [7,8]. Addressing and resolving these issues are vital not only for technological advancement but also for the tasks of ensuring high-quality medical education and protecting the rights of both practitioners and patients. Therefore, this viewpoint article aims to explore and analyze the potential ethical issues that LLMs might raise in the context of future medical education and to identify fundamental principles for developing an ethical framework for these technologies as well as corresponding methods for their implementation.

Ethical Concerns Related to the Development of LLMs

Privacy and Data Risks

In the era of electronic information technology, privacy risks and data security have always been associated with advancements in information technology. LLMs are associated with enormous numbers of parameters and training data sets and may “memorize” sensitive personal information from their training data, thus, enabling these models to generate information that contains specific individual data [9]. Jegorova et al [10] confirmed the risk of privacy leakage during the inference process exhibited by learning models. When teachers or educational institutions use data, including students' personal information, social background information, and health data, to train models to provide personalized feedback regarding students' learning interests and progress, this process entails the risk of sensitive information leakage. When LLMs are used for the purpose of assisted learning, such as in the contexts of clinical case discussions, tasks involving sensitive patient data such as name, sex, age, medical history, or final diagnoses, and even imaging data such as computed tomography scans and

magnetic resonance imaging for assistance in interpretation, patient privacy is put at risk. Previous studies have shown that even if this information is anonymized when it is input, some advanced models can reidentify such personal information from large data sets through so-called linkage attacks, thus, resulting in information exposure [10]. Rocher et al [11] reported that LLMs could use as few as 15 demographic attributes to reidentify 99.98% of personal data in any anonymized data set. Therefore, deidentifying information alone is insufficient to protect patient privacy, a situation which poses tremendous challenges with regard to protecting patient privacy when LLMs are used in the context of medical education. Furthermore, as the training and user data of LLMs are stored on cloud servers, any security vulnerabilities could result in the leakage of such sensitive information [12]. LLM development companies, such as OpenAI, may use users' personal information for service analysis, improvement, or research purposes and retain the right to share users' personal information with third parties without explicit user consent [13]. All of these factors exacerbate the risks pertaining to privacy and data security in medical education processes.

AI Hallucination

Clinical practice guidelines and consensus documents, which are continually updated, are significantly valuable as references in the context of medical education. However, “AI hallucinations” on the part of LLMs represent a major obstacle to the effective use of these documents as educational content. In the field of medicine, AI hallucinations occur when LLMs generate responses that appear to be logical but are actually incorrect, inconsistent, or fabricated, including those that rely on fabricated data and forged references [14]. Amir et al [15] compared ChatGPT 3.5 and Google BARD in terms of their ability to answer questions related to lung cancer, revealing that these LLMs exhibited error rates ranging between 17.5% and 27.5%. Similar issues have been observed in the field of dentistry, in which context LLMs can offer vague, outdated, and inaccurate information [16]. Studies have highlighted accuracy concerns with regard to LLMs' responses in medical contexts [17,18]. Several factors contribute to AI hallucinations in the field of medicine. First, LLMs' training data, which are sourced mostly from the internet, include outdated and inaccurate information that lacks quality control. Second, access to authoritative databases such as PubMed, UpToDate, and Cochrane requires subscriptions, thus, restricting LLMs from obtaining up-to-date and reliable research data [19,20]. Additionally, limitations regarding LLMs' reasoning and probabilistic text generation capabilities exacerbate the problem of AI hallucination.

Given the high level of precision required in medicine, even minor errors can have catastrophic consequences for patients. When medical educators use unverified LLM-generated content, the “one-to-many” model facilitates the widespread propagation of erroneous information. Medical students, who are in the process of acquiring foundational knowledge and trust their educators, are especially vulnerable to such misinformation. This issue can severely impact their future clinical decisions. The use of inaccurate LLM content to evaluate students reinforces incorrect information, leading to flawed assessments

and continued learning on the basis of faulty knowledge. In academic research, students may unknowingly cite false references or use incorrect data, resulting in academic misconduct. If erroneous LLM-generated information is widely disseminated in medical education, it can lead to systemic errors in clinical decision-making, thus, endangering patient health and safety. This situation undermines the rigor of medical education, erodes public trust in health care, and raises ethical issues and challenges related to academic integrity, which ultimately impact the professionalism and social responsibility of the medical field.

Training Data and Algorithmic Biases

Zack et al [21] reported that ChatGPT-4 often resorts to stereotypes and exhibits bias toward specific races, ethnicities, and sexes in terms of its diagnostic outputs. Additionally, treatment suggestions have been shown to be linked to demographic factors, to favor costlier procedures and to exhibit variations in terms of patient perceptions [21]. This widespread problem of unfair responses is primarily due to such LLMs' training data and algorithmic biases.

Bias in training data is the most direct cause of bias in LLMs' responses. LLMs learn from vast amounts of unfiltered text; thus, if these data contain biases related to sex, race, culture, or socioeconomic background, the model replicates these biases. A study on a chest x-ray classifier that had been trained mainly on White patients revealed that it lacked recognition capability in cases involving non-White patients [22]. As the amount of software-generated data increases, such information may be used as training data, further reinforcing the original biases [23].

Algorithmic bias on the part of LLMs exacerbates the problem of output bias. During the process of LLM design, the optimization goals and evaluation criteria used often focus on the majority population, neglecting the needs of the minority. For example, if an LLM is trained primarily on Western cases of alcoholic cirrhosis without considering geographic and cultural diversity or including data concerning hepatitis B-induced cirrhosis from other regions, it may provide inaccurate guidance regarding the diagnosis and treatment of hepatitis B-related cirrhosis. Failure to address training data bias and adjust algorithms perpetuates unfairness in terms of clinical practices [24]. Additionally, developers may intentionally introduce biases, which can affect data processing, model training, algorithm selection, and application, ultimately causing specific groups to receive unfair treatment. Algorithm bias impacts LLMs' responses regardless of the presence of training data bias [25].

Biases embedded in data and algorithms are reproduced in LLMs' outputs, resulting in incomplete educational content and depriving students of comprehensive and in-depth learning guidance. This lack of diversity in educational content deprives students of a balanced medical perspective, which impacts the quality and fairness of education. Future doctors may lack the knowledge and ethical sensitivity they need to serve diverse patient populations effectively, potentially leading to misdiagnoses or inappropriate treatments. Biased information based on stereotypes and unfair descriptions of specific groups can distort students' ethical concepts during the process of

diagnosis and treatment, inadvertently fostering biases against these groups. Furthermore, this gap not only affects individual students' achievements but is also amplified through medical education, thereby weakening the health care system's ability to meet the needs of diverse patient populations. LLM biases violate core biomedical ethics principles, such as nonmaleficence, beneficence, and justice. Ensuring fair and neutral LLM use in medical education is crucial with regard to efforts to maintain educational quality and uphold professional ethics.

Deficiencies in Transparency and Interpretability

When users ask questions, LLMs usually generate a predefined response. However, users often cannot comprehend how this response is generated or the underlying logic. This issue is known as the "black box" effect, in which context transparency and interpretability are lacking [26]. Transparency refers to whether the model's operation and decision-making process can be presented clearly to the user, while interpretability refers to whether users can understand the logic and reasons that underlie the decisions or text, thus, generated [9]. Such a lack of transparency and interpretability is mainly due to LLMs' large model structures and complex training data, which make their internal operations difficult to understand.

In the context of medical education, the lack of transparency and interpretability in LLMs can prevent teachers and students from assessing the accuracy of information effectively [27]. If teachers convey inaccurate information to students, such information inevitably entails risks to patient safety in clinical practice. The fact that students exhibit a high level of trust in their teachers indicates that these deficiencies can have more widespread and serious impacts than in other fields. Additionally, a lack of understanding of the logic underlying LLMs may lead to rigid thinking on the part of students, who may neglect the importance of logical thinking in medical education, thus weakening efforts to cultivate their critical thinking. Logical reasoning is crucial in clinical practice with regard to analyzing conditions accurately and formulating treatment plans. A lack of logical reasoning on the part of doctors may lead to misdiagnosis, mistreatment, and risks to patient safety. Moreover, when teachers cannot grasp the logic employed by LLMs, they may convey conclusions to students without explaining the corresponding deductive process, thus hampering their ability to guide students and decreasing the quality of teaching.

Ethical Concerns Related to the Application of LLMs

Emotional Intelligence Deficiency

Although LLMs are trained to provide empathetic responses in the context of engaging with patients [28]—recent research has shown that ChatGPT even exhibits higher levels of empathy when consulting on systemic sclerosis than do neurologists [29]—the artificial empathy expressed by AI generally cannot replicate the subtle genuineness of the emotions conveyed by health care professionals [30]. True emotional expression is the result of perceiving and understanding the patient's emotions through real-time communication, especially by sensing changes in tone, facial expressions, and body language, with the goals

of comprehending the patient's needs and concerns and providing personalized feedback. However, existing LLMs cannot perceive these nonverbal emotional changes. Furthermore, the expression of empathy by LLMs is limited to textual descriptions of emotions, whereas variations in tone, facial expressions, and body language on the part of health care professionals, such as handshakes or hugs, often convey empathy more effectively than can words alone [1]. Prof Marc Succi suggested that excessive reliance on artificial emotional communication in medical settings may exacerbate the societal public health issue of "loneliness" [31]. Establishing and developing empathy involves addressing the complexity and uncertainty that characterize real interactions, which are shaped by patients' diverse backgrounds, beliefs, and values. LLMs, as teaching tools, cannot convey a diverse range of emotional expressions beyond the level permitted by text, thus, preventing medical students from gathering materials that can help them improve their communication skills and empathy through such interactions. In addition to direct patient communication, teachers who exhibit empathy and care in clinical settings also serve as good examples who can enhance students' empathy. Sole reliance on LLMs for tasks such as medical history collection, condition disclosure, and obtaining consent can reduce the presence of empathy in clinical environments, thus, conflicting with the goals of medical ethics education. Excessive reliance on artificial emotional intelligence can lead to a lack of empathy among medical students. The development of empathy requires ongoing practice in diverse emotional settings, which static LLMs cannot provide.

Educational Unfairness

Unfairness in the context of traditional medical education is characterized primarily by several aspects: uneven resource distribution, disparities in faculty strength, high tuition fees that limit student access, and geographical disparities in terms of school location [32]. While LLMs offer rich educational resources and opportunities, their application is a double-edged sword with respect to educational fairness.

For institutions that have only limited teaching resources, LLMs can aggregate rich teaching resources at a global level, thus, providing students with high-quality educational experiences. They also offer relatively equal learning opportunities to regions located in remote geographical areas and enable nonnative speakers to access and understand medical knowledge in other languages through cross-language support [33]. However, medical schools that have better technological infrastructure, higher budgets, and access to the most up-to-date technologies may find it easier to take advantage of the benefits of LLMs in the context of education. In contrast, schools and regions with limited resources may experience exacerbated resource allocation inequalities because of their inability to afford these technologies, their lack of the professional knowledge necessary to implement these technologies, and restrictions on accessing LLMs (such as ChatGPT, which is currently unavailable to Chinese users). Moreover, the use of LLMs requires extensive data processing, which may expose data to privacy and security risks. Institutions that lack sufficient resources and robust infrastructure may struggle to safeguard user data, thus, increasing the risk of data leakage and misuse. With respect to

academic achievements, individuals who can use these tools may gain access to more recognition and opportunities within the academic community, whereas students and researchers who lack such access may face disadvantages in terms of academic competition [34].

Academic Integrity Concerns

Issues related to academic integrity in the context of medical education notably include plagiarism and cheating. LLMs greatly facilitate medical research by summarizing published articles and assisting with data analysis or literature retrieval. Research has indicated that when abstracts generated by LLMs and those authored by humans are presented simultaneously, university professors struggle to distinguish between them [35]. Another study that examined LLMs' generation of fictitious article abstracts revealed that peer reviewers could identify only 68% of abstracts generated by ChatGPT as fictitious [36]. This outcome indicates that the quality of the text produced by LLMs has approached or reached a professional level, thus, giving rise to opportunities for academic plagiarism; namely, students may use LLMs to generate abstracts, data, or even complete draft papers. Despite the originality of the content generated by LLMs, students who submit these texts as their own work may be suspected of plagiarism. With regard to one recently published article, the direct replication of generated content, including the prompts from LLMs—the regenerated response—resulted in article retraction [37]. Numerous other articles now face similar scrutiny, and the question of whether these will be retracted by the journal or revised remains undecided [38,39]. This excessive reliance and lack of critical use of LLMs have caused serious damage to academic integrity and posed ethical challenges. Furthermore, the integration of LLMs into browsers facilitates direct searches for source texts or other multimedia documents, which can exacerbate issues related to plagiarism and make them more prevalent when teachers or students use these data directly. A well-known test of the accuracy of LLMs in the context of medical questions is ChatGPT's successful passing of the United States Medical Licensing Exam [40]. Many studies have shown that LLMs achieve high scores in various disciplines [41-43]. This high level of accuracy enables students to use LLMs to cheat on their examinations. Excessive reliance on LLMs may decrease students' research ability, critical thinking, innovation, and motivation to engage in proper learning [44]. This situation can perpetuate students' reliance on LLMs and academic misconduct, thus, leading to a vicious cycle. Additionally, if students and teachers plagiarize LLM-generated content without critical evaluation, erroneous or biased information may become widespread.

Responsibility and Copyright Concerns

In the context of traditional clinical education, clinical teachers supervise students' operations, provide students with guidance, and take responsibility for adverse clinical events, while students adhere to medical standards and ethical norms. Medical institutions must ensure a safe learning environment and sufficient supervision. This clear division of responsibilities can help protect patient safety and improve the quality of education.

In the future, the introduction of LLMs into medical education will inevitably add new dimensions to the division of responsibilities. From a legal perspective, AI lacks the legal standing of a human being, thus leaving humans as the ultimate accountable parties [45]. However, in medical settings, the responsibility for patient harm due to LLM biases and inaccuracies is unclear, a situation which affects doctors, patients, and institutions such as medical facilities and developers. LLM developers should ensure that their tools are accurate and safe, but entities such as OpenAI disclaim responsibility for the texts generated by their LLMs [46]. Existing ethical and legal frameworks have not fully adapted to the challenges posed by such emerging tools, thus, leading to ambiguity in terms of responsibility attribution and regulation in the context of actual application.

When teachers use LLMs to generate teaching outlines, lecture notes, or textbooks or when students use LLMs to write academic papers—especially when LLMs access original texts and images from the internet through search engines—and then copy from copyrighted source materials directly or borrow from them excessively, copyright disputes can occur [4]. To protect copyrights and maintain academic integrity more effectively, current guidelines issued by the International Committee of Medical Journal Editors and the Committee on Public Ethics suggest that LLMs such as ChatGPT should not be listed as authors of papers [47]. LLMs cannot bear the same responsibility as human authors or provide substantive explanations for the content they generate. However, this approach is not universally accepted, and the role of LLMs in authorship declarations remains the subject of ongoing discussion by various publishing institutions [48,49].

AI-Related Laws and Ethics

Overview

Due to the advancement of AI technology in various domains, particularly with regard to the development of LLMs, multiple countries and international organizations have actively formulated or updated ethical and legal guidelines. These measures focus on key areas such as privacy protection, transparency, algorithmic fairness, and accountability and offered reliable guidance concerning the reasonable use and development of artificial intelligence technologies. AI legal and ethical principles are in [Multimedia Appendix 1](#).

Laws

The General Data Protection Regulation of the European Union, which was enacted in 2018, offers stringent guidelines for handling sensitive data in the education and health care sectors. These guidelines emphasize the pseudonymization or anonymization of personal data storage, requiring the highest privacy settings, and ensuring transparency in terms of data processing. Additionally, they strengthen users' control over their personal information, thus, empowering individuals to manage their data effectively [50]. In China, the Personal Information Protection Law, which was implemented in 2021, establishes rigorous standards for the protection of personal information, particularly with regard to user awareness and consent rights [51]. Similarly, the Health Insurance Portability

and Accountability Act in the United States prioritizes the protection of medical information and respect for patient privacy [52].

Ethical Guidelines and Initiatives

International guiding principles have extensively discussed issues pertaining to transparency, interpretability, algorithmic fairness, and accountability in the context of AI. The Institute of Electrical and Electronics Engineers Global Initiative on Ethics of Autonomous and Intelligent Systems, which was founded in 2019, emphasized the transparency of autonomous systems in AI design processes and ethical guidelines for handling private data and algorithmic biases [53]. The European Union's White Paper on Artificial Intelligence, which was released in February 2020, elaborated on ethical frameworks for system transparency, privacy data, and algorithmic biases, particularly with regard to the importance of avoiding bias and discrimination and ensuring the fair use of AI decision support systems across diverse populations [54]. The United Nations Educational, Scientific and Cultural Organization's "Recommendations on the Ethics of Artificial Intelligence," which was issued in November 2021, systematically addressed privacy rights, data protection, transparency, interpretability, responsibility, and accountability in the context of AI applications, emphasizing fairness and nondiscrimination [55]. In January 2024, the World Health Organization issued the "Ethics and Governance of AI in Health: Guidance on Large Multimodal Models," which was the first ethical standard for large multimodal models such as LLMs. This document emphasized transparency, interpretability, and accountability; promoted fair access; and prioritized inclusivity [56]. In March 2024, the United Nations General Assembly adopted a US-led resolution titled "Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development," which represented the first global consensus on AI governance. This resolution addressed the needs to bridge regional AI gaps, protect against discrimination and bias, safeguard privacy, and respect intellectual property (IP) rights, among other 13 proposals [57].

Limitations

Existing legal and ethical guidelines ensure the safe, transparent, and fair use of AI systems. However, when LLMs are used in the context of medical education, significant limitations emerge. First, these guidelines, which have been shaped by various national, organizational, and cultural contexts, reflect different social values, historical backgrounds, and priorities. For instance, the General Data Protection Regulation and the Personal Information Protection Law in China emphasize data privacy, whereas the Institute of Electrical and Electronics Engineers guidelines prioritize technological innovation. Such differences can lead to conflicts and increase the complexity and uncertainty of applying LLMs across different countries. Second, despite the fact that these guidelines emphasize transparency and interpretability, the complex algorithms and nonlinear structures of LLMs make these principles difficult to achieve. Existing guidelines lack detailed, actionable recommendations that can help users understand and trust LLM decision-making in the context of medical education.

Furthermore, LLMs may amplify biases in their training data, thus, leading to unfair outcomes. While many guidelines have emphasized the importance of avoiding discrimination and bias, they have not provided specific tools or methods to detect and correct biases within LLMs. Additionally, current ethical oversight and accountability mechanisms are inadequate with respect to the use of LLMs in medical education. When LLMs generate misleading or inaccurate content that could cause harm, existing frameworks lack clear guidelines for determining when human intervention is necessary, how to monitor and adjust the model's output, and how to establish accountability. Most importantly, existing legal and ethical guidelines in the field of medical education are fragmented and lack cohesiveness. They do not comprehensively address unique challenges such as academic misconduct, plagiarism, and copyright issues. Consequently, educators often struggle to find clear, authoritative solutions in practice. Therefore, it is critical to

develop an ethical framework that is specifically tailored to the application of LLMs in medical education, thereby ensuring the safety, efficacy, and fairness of such technologies and promoting their widespread and in-depth application in this critical field.

Fundamental Principles of and Ways of Implementing LLM Ethics in Medical Education

Overview

To ensure that LLMs are used within appropriate legal and ethical frameworks, it is essential to establish a unified ethical framework. In light of the broad biomedical principles proposed by Beauchamp and Childress, namely, beneficence, justice, autonomy, and nonmaleficence [5], we propose fundamental principles and implementation methods for the use of LLMs in the context of medical education that are in line with these biomedical ethical principles (Figure 1 and Table 1).

Figure 1. Eight fundamental principles regarding the use of large language models in medical education.



Table 1. Implementation methods aligned with biomedical ethical principles regarding the use of large language models (LLMs) in medical education.

Ethical principles for LLMs in medical education	Implementation methods	Biomedical ethical principles
Privacy and data protection	<ol style="list-style-type: none"> 1. Multilayered privacy protection mechanisms 2. Data encryption and secure storage 3. Strict data usage policies 4. Strict access controls 	Nonmaleficence and autonomy
Data quality control and supervision	<ol style="list-style-type: none"> 1. Develop independent LLM databases 2. Link to authoritative databases 3. Maintain a critical approach 	Beneficence and justice
Fairness and equal treatment	<ol style="list-style-type: none"> 1. Diverse and comprehensive training data 2. Fairness review and monitoring mechanisms 3. Open and accessible resources and tools 4. Cross-regional and cross-economic background collaboration 	Justice
Transparency and interpretability	<ol style="list-style-type: none"> 1. Model documentation and operational guidelines 2. Implement interpretative interfaces or tools 3. Regular public evaluations and feedback 	Autonomy
Academic integrity and moral norms	<ol style="list-style-type: none"> 1. Establish policies and guidelines 2. Develop and use detection tools 3. Maintain educators' leadership roles 4. Foster empathy and compassion 	Nonmaleficence and beneficence
Accountability and traceability	<ol style="list-style-type: none"> 1. Clear responsibility framework 2. Detailed logging and monitoring systems 3. Independent review and investigation committees 4. Clear complaint and feedback channels 	Nonmaleficence and justice
Intellectual property protection	<ol style="list-style-type: none"> 1. Strict IP compliance processes 2. Copyright statements and terms of use 3. Monitoring and detection systems 4. Complaint and feedback mechanisms 	Justice and autonomy
Educational research and innovation	<ol style="list-style-type: none"> 1. Establish dedicated research and development funds 2. Creating an open teaching experimentation environment 3. Optimizing teaching content 	Beneficence

Privacy and Data Protection

To mitigate the risk of sensitive data breaches effectively and ensure both data security and privacy protection, it is crucial to implement a series of robust data protection measures involving dedicated LLM training databases that are specifically designed for medical education.

Multilayered Privacy Protection Mechanisms: Advanced data deidentification techniques, such as generalizing and suppressing sensitive information, must be used to prevent reidentification. Differential privacy techniques must be employed to enhance protection by adding noise to the data [58].

Data Encryption and Secure Storage: Data encryption and storage must be enhanced based on tamper-proof hardware security modules and distributed storage systems with the goals of preventing attacks and ensuring data security in terms of both storage and transmission [59].

Strict Data Usage Policies: Strict policies must be established and followed to ensure that data providers have control over and transparency regarding their data use. Users must be

informed about data usage and allowed to withdraw their consent at any time.

Strict Access Controls: Multifactor authentication and stringent access controls, such as SMS text messaging verification and facial recognition, must be implemented to ensure that only authorized personnel access sensitive data [60].

Data Quality Control and Supervision

To prevent partial and inaccurate responses based on erroneous or outdated training data, effective quality control and supervision mechanisms must be established.

Develop independent LLM databases: all training data should be reviewed by clinical experts before incorporation to ensure authenticity and accuracy. LLM-generated content must be regularly reviewed and evaluated to identify and correct biases and errors promptly.

Link to authoritative databases: authoritative databases such as PubMed and UpToDate must be used to ensure the accuracy and relevance of training data. These resources are to be updated regularly with the latest medical research, thereby reducing errors and biases as well as enhancing data credibility.

Maintain a critical approach: educators and students should assess LLM-generated information critically by verifying it by reference to authoritative databases and other language models. This practice can ensure high-quality standards in medical education involving LLMs.

Fairness and Equal Treatment

Mitigating biases in LLMs can ensure that all students and educators have equal opportunities and resources in the context of medical education, thereby eliminating the health disparities and knowledge inequalities caused by these biases.

Diverse and comprehensive training data: developers must use diverse datasets during model training to avoid biases. This process involves reviewing and filtering data with the goal of eliminating the factors that lead to discrimination or unfair treatment.

Fairness review and monitoring mechanisms: fairness review and monitoring mechanisms must be incorporated into the model's deployment and usage. The biases or unjust tendencies exhibited by the model outputs must be evaluated and detected regularly, and adjustments and optimizations must be made as needed.

Open and accessible resources and tools: open and accessible resources and tools must be provided to students and educators at various economic levels and geographic locations to enable them to use LLMs on an equal basis. For example, developing lightweight or localized versions of a model can support advanced technology adoption in resource-limited areas.

Cross-regional and cross-economic background collaboration: cross-regional and cross-economic collaboration platforms and communities must be established and supported with the goal of promoting the sharing of knowledge and resources, thereby mitigating the educational inequalities caused by regional disparities.

Transparency and Interpretability

Ensuring that the operations and decision-making processes of LLMs are transparent and explainable to students and educators is critical with respect to enhancing students' and educators' understanding and trust in the system.

Model documentation and operational guidelines: developers should provide comprehensive documentation and guidelines to explain the model's mechanisms, input-output processes, and decision-making logic. These guidelines should cover design objectives, the sources of training data and corresponding methods, and potential limitations with the aim of helping users understand the model's construction and functionality.

Interpretative interfaces or tools: in the context of model use, interpretative interfaces or tools must be implemented to allow students and educators to understand the model's reasoning and decisions in real time. For example, visual tools can reveal how the model processes input data and generate outputs, thereby enhancing the model's transparency and interpretability [61].

Regular public evaluations and feedback: involving experts and users in regular evaluations and reviews and establishing effective feedback mechanisms can help identify and correct

issues or biases associated with the use of the model, thereby promoting continuous improvement and optimization.

Academic Integrity and Moral Norms

An emphasis on academic integrity and ethical norms can prevent students and educators from gaining undue advantages from LLMs or from engaging in misconduct, such as plagiarism or cheating. Students and educators should use LLMs reasonably, avoid excessive reliance on such technologies, and focus on cultivating empathy, compassion, and critical thinking.

Establish policies and guidelines: educational institutions should establish specific LLM usage policies and guidelines based on a unified set of ethical principles to specify the legal and ethical use of LLMs as well as to prohibit any form of academic misconduct clearly and explain the corresponding consequences.

Develop and use detection tools: detection tools must be developed to identify and prevent inappropriate content or plagiarism with regard to LLM-generated material, thus ensuring that students' submissions and research results are genuine and original [62].

Maintain educators' leadership roles: during the teaching process, educators should continue to play a leadership role by encouraging students to analyze and evaluate LLM-generated content critically and emphasizing that models are auxiliary tools rather than substitutes for human thought. Educators should encourage students to question and verify information from LLMs, thereby enhancing their independent thinking skills.

Foster empathy and compassion: in clinical practice, cultivating students' empathy and compassion through rich interpersonal interactions can enhance the humanistic aspect of patient care. Educators should lead by example, such as by demonstrating material and guiding students in the process of applying these skills in real clinical settings. This approach can help students understand and respect patients' emotions and needs, ultimately improving the overall quality of health care services.

Accountability and Traceability

To ensure that responsibilities and accountability procedures are defined clearly when LLM decision errors negatively affect patient health, clear accountability and traceability mechanisms must be established.

Clear framework for responsibility: developing a clear framework of responsibilities and obligations for model developers, deployers, and users at different stages is crucial. Developers should ensure that the model's design and training meet safety and fairness standards; deployers must ensure compliance with supervision and quality control measures; and users must make and verify professional judgments on the basis of the model's recommendations.

Detailed logging and monitoring systems: implementing detailed logging and monitoring during the use of LLMs involves recording the model's inputs, outputs, and decision paths, which can make it possible to trace the decision-making process when issues arise.

Independent review and investigation committees: independent committees must be established to address and investigate health

issues resulting from LLM decision errors while ensuring fair procedures. These committees should include technical experts, legal advisors, and medical professionals who can assess responsibilities jointly and recommend improvements.

Clear complaint and feedback channels: providing clear complaint and feedback channels for patients and users allows them to report issues and seek help promptly in cases involving harm.

Intellectual Property Protection

Ensuring compliance with IP laws during the use of LLMs and protecting the rights of knowledge creators and holders are critical to the task of preventing unauthorized knowledge use and infringement.

Strict IP compliance processes: strict IP compliance processes must be established to ensure that LLMs use only authorized or publicly licensed data. Developers should document data sources meticulously and review them regularly for legality and compliance.

Copyright statements and terms of use: when LLM-generated content is used, clear copyright statements and terms of use must be provided. Such statements and terms of use can ensure that users understand and comply with IP laws, thus preventing the unauthorized creation or dissemination of protected content.

Monitoring and detection systems: effective monitoring and detection systems can be implemented to identify and prevent unauthorized use and potential infringement. For example, copyright detection tools can automatically flag potentially infringing content in model outputs [63].

Complaint and feedback mechanisms: complaint and feedback mechanisms can be established to allow creators and IP holders to report and address potential infringement issues promptly, thus facilitating efficient remedial actions.

Educational Research and Innovation

Taking advantage of the benefits of LLM technology for educational research and innovation can enhance the quality and effectiveness of medical education, thus promoting the continuous improvement and optimization of teaching methods and content.

Establish dedicated research and development funds: dedicated research and development funds must be established to support the attempts of educational institutions and academic groups to conduct LLM-centered innovative educational projects and research. Encouraging interdisciplinary collaboration and integrating the most recent advancements in AI and medical education can facilitate the exploration of novel teaching methods and content.

Open teaching experimentation environment: establishing an open experimentation and testing environment can allow educators to explore innovative teaching methods safely. The powerful analytical capabilities of LLMs can be used to evaluate the effectiveness of these methods. The development of LLM-based intelligent tutoring tools and digital learning

modules can further optimize teaching strategies and enhance the student learning experience, ultimately improving overall teaching quality.

Optimizing teaching content: to ensure that medical education remains aligned with the most up-to-date advancements in medical research and technology, researchers should regularly monitor and evaluate recent findings and clinical practices. These insights must be integrated into the curriculum via LLMs to ensure that the curriculum remains current and promotes continuous improvement.

While we have to identify certain fundamental principles and methods for an ethical framework for LLMs in medical education, several key areas require further refinement. First, this framework should be dynamic and adaptable to facilitate continuous updates as technology and society evolve. Second, a cross-cultural and global perspective is crucial with regard to respecting and accommodating ethical norms associated with different backgrounds. Additionally, involving a wide range of stakeholders—data providers, developers, policymakers, educators, students, and patients—is vital if their needs and concerns are to be reflected. Finally, strengthening ethical education and training is essential for stakeholders' ability to understand and apply these principles. In the future, adhering to a unified ethical framework can establish a safe, transparent, fair, and high-quality environment to support the development of medical education.

Conclusions

This paper explores the complex challenges that LLMs may entail with regard to the future of medical education, and it covers the entire spectrum ranging from the development of such technologies to their practical application. It examines the limitations of current AI-related legal and ethical frameworks with regard to guiding LLM use in the context of medical education and advocates for the development of a unified ethical framework specifically for this purpose. This paper highlights 8 fundamental principles and detailed implementation measures necessary to create such a framework: privacy and data protection, data quality control and supervision, fairness and equal treatment, transparency and interpretability, academic integrity and moral norms, accountability and traceability, IP protection, and educational research and innovation. These principles provide a solid foundation for the development of a comprehensive and actionable ethical framework. Adhering to this unified ethical framework can offer clear guidance concerning the application of LLMs in medical education. Future research should refine and implement this ethical framework in the context of a rapidly evolving technological landscape, including by facilitating dynamic updates in light of new technologies, ensuring global cross-cultural applicability, engaging diverse stakeholders, and strengthening ethical education and training. These efforts aim to balance technological advancement with ethical values, thus, promoting progress in medical education without compromising the principles of fairness, equity, or patient safety with the goal of ensuring higher quality and greater inclusivity.

Acknowledgments

We used the cross-language capabilities of ChatGPT-4 and ChatGPT-4o to enhance the language used in this paper. The revisions were reviewed by all the authors and subsequently refined and polished by American Journal Experts to ensure publication quality. We thank Ms Ding Ding for her assistance in the literature search.

Authors' Contributions

LZ and RW contributed to the conceptual design, data collection, drafting, final revision, and final submission of this paper. LFH, WXH, and FQN contributed to the data collection, language correction, and revision of this paper. All the authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Artificial intelligence–related legal and ethical principles.

[\[DOCX File , 18 KB-Multimedia Appendix 1\]](#)

References

1. 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;9:e50945. [[FREE Full text](#)] [doi: [10.2196/50945](https://doi.org/10.2196/50945)] [Medline: [37578830](https://pubmed.ncbi.nlm.nih.gov/37578830/)]
2. Xu X, Wang X, Zhang Y, Zheng R. Applying ChatGPT to tackle the side effects of personal learning environments from learner and learning perspective: an interview of experts in higher education. *PLoS One*. 2024;19(1):e0295646. [[FREE Full text](#)] [doi: [10.1371/journal.pone.0295646](https://doi.org/10.1371/journal.pone.0295646)] [Medline: [38170691](https://pubmed.ncbi.nlm.nih.gov/38170691/)]
3. Kavadella A, Dias da Silva MA, Kaklamanos E, Stamatopoulos V, Giannakopoulos K. Evaluation of ChatGPT's real-life implementation in undergraduate dental education: mixed methods study. *JMIR Med Educ*. 2023;10:e51344. [[FREE Full text](#)] [doi: [10.2196/51344](https://doi.org/10.2196/51344)] [Medline: [38111256](https://pubmed.ncbi.nlm.nih.gov/38111256/)]
4. Abd-Alrazaq A, AlSaad R, Alhuwail D, Ahmed A, Healy PM, Latifi S, et al. Large language models in medical education: opportunities, challenges, and future directions. *JMIR Med Educ*. 2023;9:e48291. [[FREE Full text](#)] [doi: [10.2196/48291](https://doi.org/10.2196/48291)] [Medline: [37261894](https://pubmed.ncbi.nlm.nih.gov/37261894/)]
5. Beauchamp T, Childress J. Principles of biomedical ethics: marking its fortieth anniversary. *Am J Bioeth*. 2019;19(11):9-12. [doi: [10.1080/15265161.2019.1665402](https://doi.org/10.1080/15265161.2019.1665402)] [Medline: [31647760](https://pubmed.ncbi.nlm.nih.gov/31647760/)]
6. Busch F, Adams LC, Bressemer KK. Biomedical ethical aspects towards the implementation of artificial intelligence in medical education. *Med Sci Educ*. 2023;33(4):1007-1012. [[FREE Full text](#)] [doi: [10.1007/s40670-023-01815-x](https://doi.org/10.1007/s40670-023-01815-x)] [Medline: [37546190](https://pubmed.ncbi.nlm.nih.gov/37546190/)]
7. Çalışkan SA, Demir K, Karaca O. Artificial intelligence in medical education curriculum: an e-Delphi study for competencies. *PLoS One*. 2022;17(7):e0271872. [[FREE Full text](#)] [doi: [10.1371/journal.pone.0271872](https://doi.org/10.1371/journal.pone.0271872)] [Medline: [35862401](https://pubmed.ncbi.nlm.nih.gov/35862401/)]
8. Charow R, Jeyakumar T, Younus S, Dolatabadi E, Salhia M, Al-Mouaswas D, et al. Artificial intelligence education programs for health care professionals: scoping review. *JMIR Med Educ*. 2021;7(4):e31043. [[FREE Full text](#)] [doi: [10.2196/31043](https://doi.org/10.2196/31043)] [Medline: [34898458](https://pubmed.ncbi.nlm.nih.gov/34898458/)]
9. Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical considerations of using ChatGPT in health care. *J Med Internet Res*. 2023;25:e48009. [[FREE Full text](#)] [doi: [10.2196/48009](https://doi.org/10.2196/48009)] [Medline: [37566454](https://pubmed.ncbi.nlm.nih.gov/37566454/)]
10. Jegorova M, Kaul C, Mayor C, O'Neil AQ, Weir A, Murray-Smith R, et al. Survey: leakage and privacy at inference time. *IEEE Trans Pattern Anal Mach Intell*. 2023;45(7):9090-9108. [[FREE Full text](#)] [doi: [10.1109/TPAMI.2022.3229593](https://doi.org/10.1109/TPAMI.2022.3229593)] [Medline: [37015684](https://pubmed.ncbi.nlm.nih.gov/37015684/)]
11. Rocher L, Hendrickx JM, de Montjoye Y. Estimating the success of re-identifications in incomplete datasets using generative models. *Nat Commun*. 2019;10(1):3069. [[FREE Full text](#)] [doi: [10.1038/s41467-019-10933-3](https://doi.org/10.1038/s41467-019-10933-3)] [Medline: [31337762](https://pubmed.ncbi.nlm.nih.gov/31337762/)]
12. Li J. Security implications of AI Chatbots in health care. *J Med Internet Res*. 2023;25:e47551. [[FREE Full text](#)] [doi: [10.2196/47551](https://doi.org/10.2196/47551)] [Medline: [38015597](https://pubmed.ncbi.nlm.nih.gov/38015597/)]
13. Data Usage for Consumer Services FAQ. URL: <https://help.openai.com/en/articles/7039943-data-usage-for-consumer-services-faq> [accessed 2024-04-18]
14. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*. 2023;15(2):e35179. [[FREE Full text](#)] [doi: [10.7759/cureus.35179](https://doi.org/10.7759/cureus.35179)] [Medline: [36811129](https://pubmed.ncbi.nlm.nih.gov/36811129/)]
15. Rahsepar AA, Tavakoli N, Kim GHJ, Hassani C, Abtin F, Bedayat A. How AI responds to common lung cancer questions: ChatGPT vs Google Bard. *Radiology*. 2023;307(5):e230922. [doi: [10.1148/radiol.230922](https://doi.org/10.1148/radiol.230922)] [Medline: [37310252](https://pubmed.ncbi.nlm.nih.gov/37310252/)]
16. Giannakopoulos K, Kavadella A, Aaqel Salim A, Stamatopoulos V, Kaklamanos EG. Evaluation of the performance of generative AI large language models ChatGPT, Google Bard, and Microsoft Bing Chat in supporting evidence-based

- dentistry: comparative mixed methods study. *J Med Internet Res.* 2023;25:e51580. [FREE Full text] [doi: [10.2196/51580](https://doi.org/10.2196/51580)] [Medline: [38009003](https://pubmed.ncbi.nlm.nih.gov/38009003/)]
17. Zúñiga Salazar G, Zúñiga D, Vindel CL, Yoong AM, Hincapie S, Zúñiga B, et al. Efficacy of AI chats to determine an emergency: a comparison between OpenAI's ChatGPT, Google Bard, and Microsoft Bing AI Chat. *Cureus.* 2023;15(9):e45473. [FREE Full text] [doi: [10.7759/cureus.45473](https://doi.org/10.7759/cureus.45473)] [Medline: [37727841](https://pubmed.ncbi.nlm.nih.gov/37727841/)]
 18. Lim ZW, Pushpanathan K, Yew SME, Lai Y, Sun C, Lam JSH, et al. Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard. *EBioMedicine.* 2023;95:104770. [FREE Full text] [doi: [10.1016/j.ebiom.2023.104770](https://doi.org/10.1016/j.ebiom.2023.104770)] [Medline: [37625267](https://pubmed.ncbi.nlm.nih.gov/37625267/)]
 19. Xue E, Bracken-Clarke D, Iannantuono GM, Choo-Wosoba H, Gulley JL, Floudas CS. Utility of large language models for health care professionals and patients in navigating hematopoietic stem cell transplantation: comparison of the performance of ChatGPT-3.5, ChatGPT-4, and Bard. *J Med Internet Res.* 2024;26:e54758. [FREE Full text] [doi: [10.2196/54758](https://doi.org/10.2196/54758)] [Medline: [38758582](https://pubmed.ncbi.nlm.nih.gov/38758582/)]
 20. Arif TB, Munaf U, Ul-Haque I. The future of medical education and research: is ChatGPT a blessing or blight in disguise? *Med Educ Online.* 2023;28(1):2181052. [FREE Full text] [doi: [10.1080/10872981.2023.2181052](https://doi.org/10.1080/10872981.2023.2181052)] [Medline: [36809073](https://pubmed.ncbi.nlm.nih.gov/36809073/)]
 21. Zack T, Lehman E, Suzgun M, Rodriguez JA, Celi LA, Gichoya J, et al. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. *Lancet Digit Health.* 2024;6(1):e12-e22. [FREE Full text] [doi: [10.1016/S2589-7500\(23\)00225-X](https://doi.org/10.1016/S2589-7500(23)00225-X)] [Medline: [38123252](https://pubmed.ncbi.nlm.nih.gov/38123252/)]
 22. Seyyed-Kalantari L, Liu G, McDermott M, Chen IY, Ghassemi M. CheXclusion: fairness gaps in deep chest X-ray classifiers. *Pac Symp Biocomput.* 2021;26:232-243. [FREE Full text] [Medline: [33691020](https://pubmed.ncbi.nlm.nih.gov/33691020/)]
 23. Akinci D'Antonoli T, Stanzione A, Bluethgen C, Vernuccio F, Ugga L, Klontzas ME, et al. Large language models in radiology: fundamentals, applications, ethical considerations, risks, and future directions. *Diagn Interv Radiol.* 2024;30(2):80-90. [FREE Full text] [doi: [10.4274/dir.2023.232417](https://doi.org/10.4274/dir.2023.232417)] [Medline: [37789676](https://pubmed.ncbi.nlm.nih.gov/37789676/)]
 24. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Large language models encode clinical knowledge. *Nature.* 2023;620(7972):172-180. [FREE Full text] [doi: [10.1038/s41586-023-06291-2](https://doi.org/10.1038/s41586-023-06291-2)] [Medline: [37438534](https://pubmed.ncbi.nlm.nih.gov/37438534/)]
 25. Khera R, Butte AJ, Berkwitz M, Hswen Y, Flanagan A, Park H, et al. AI in Medicine-JAMA's focus on clinical outcomes, patient-centered care, quality, and equity. *JAMA.* 2023;330(9):818-820. [doi: [10.1001/jama.2023.15481](https://doi.org/10.1001/jama.2023.15481)] [Medline: [37566406](https://pubmed.ncbi.nlm.nih.gov/37566406/)]
 26. Stokel-Walker C, Van Noorden R. What ChatGPT and generative AI mean for science. *Nature.* 2023;614(7947):214-216. [doi: [10.1038/d41586-023-00340-6](https://doi.org/10.1038/d41586-023-00340-6)] [Medline: [36747115](https://pubmed.ncbi.nlm.nih.gov/36747115/)]
 27. Patino GA, Amiel JM, Brown M, Lypson ML, Chan TM. The promise and perils of artificial intelligence in health professions education practice and scholarship. *Acad Med.* 2024;99(5):477-481. [doi: [10.1097/ACM.0000000000005636](https://doi.org/10.1097/ACM.0000000000005636)] [Medline: [38266214](https://pubmed.ncbi.nlm.nih.gov/38266214/)]
 28. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence Chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med.* 2023;183(6):589-596. [FREE Full text] [doi: [10.1001/jamainternmed.2023.1838](https://doi.org/10.1001/jamainternmed.2023.1838)] [Medline: [37115527](https://pubmed.ncbi.nlm.nih.gov/37115527/)]
 29. Maida E, Moccia M, Palladino R, Borriello G, Affinito G, Clerico M, et al. Digital Technologies, Web, Social Media Study Group of the Italian Society of Neurology (SIN). ChatGPT vs. neurologists: a cross-sectional study investigating preference, satisfaction ratings and perceived empathy in responses among people living with multiple sclerosis. *J Neurol.* 2024;271(7):4057-4066. [doi: [10.1007/s00415-024-12328-x](https://doi.org/10.1007/s00415-024-12328-x)] [Medline: [38568227](https://pubmed.ncbi.nlm.nih.gov/38568227/)]
 30. Guidi C, Traversa C. Empathy in patient care: from 'clinical empathy' to 'empathic concern'. *Med Health Care Philos.* 2021;24(4):573-585. [FREE Full text] [doi: [10.1007/s11019-021-10033-4](https://doi.org/10.1007/s11019-021-10033-4)] [Medline: [34196934](https://pubmed.ncbi.nlm.nih.gov/34196934/)]
 31. Koranteng E, Rao A, Flores E, Lev M, Landman A, Dreyer K, et al. Empathy and equity: key considerations for large language model adoption in health care. *JMIR Med Educ.* 2023;9:e51199. [FREE Full text] [doi: [10.2196/51199](https://doi.org/10.2196/51199)] [Medline: [38153778](https://pubmed.ncbi.nlm.nih.gov/38153778/)]
 32. Tsouroufli M, Malcolm I. Equality, diversity and fairness in medical education: international perspectives. *Med Educ.* 2015;49(1):4-6. [doi: [10.1111/medu.12601](https://doi.org/10.1111/medu.12601)] [Medline: [25545562](https://pubmed.ncbi.nlm.nih.gov/25545562/)]
 33. Biswas S. ChatGPT and the future of medical writing. *Radiology.* 2023;307(2):e223312. [doi: [10.1148/radiol.223312](https://doi.org/10.1148/radiol.223312)] [Medline: [36728748](https://pubmed.ncbi.nlm.nih.gov/36728748/)]
 34. Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Front Artif Intell.* 2023;6:1169595. [FREE Full text] [doi: [10.3389/frai.2023.1169595](https://doi.org/10.3389/frai.2023.1169595)] [Medline: [37215063](https://pubmed.ncbi.nlm.nih.gov/37215063/)]
 35. Else H. Abstracts written by ChatGPT fool scientists. *Nature.* 2023;613(7944):423. [doi: [10.1038/d41586-023-00056-7](https://doi.org/10.1038/d41586-023-00056-7)] [Medline: [36635510](https://pubmed.ncbi.nlm.nih.gov/36635510/)]
 36. Gao CA, Howard FM, Dyer EC, Ramesh S, Luo Y. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digit Med.* 2023;6(1):75. [doi: [10.1038/s41746-023-00819-6](https://doi.org/10.1038/s41746-023-00819-6)] [Medline: [37100871](https://pubmed.ncbi.nlm.nih.gov/37100871/)]
 37. Sibel T, Karmina K, Abdullahi Y. Retracted: exploring new optical solutions for nonlinear Hamiltonian amplitude equation via two integration schemes. *Phys.* 2023. URL: <https://iopscience.iop.org/article/10.1088/1402-4896/aceb40/meta> [accessed 2024-04-30]

38. Zhang M, Wu L, Yang T, Zhu B, Liu Y. The three-dimensional porous mesh structure of Cu-based metal-organic-framework - aramid cellulose separator enhances the electrochemical performance of lithium metal anode batteries. *Surfaces and Interfaces*. 2024;46:104081. [doi: [10.1016/j.surfin.2024.104081](https://doi.org/10.1016/j.surfin.2024.104081)]
39. Patil M, Sidramappa A, Shetty SK, Hebbale AM. Experimental study of solar PV/T panel to increase the energy conversion efficiency by air cooling. *Materials Today: Proceedings*. 2023;92:309-313. [doi: [10.1016/j.matpr.2023.05.007](https://doi.org/10.1016/j.matpr.2023.05.007)]
40. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health*. 2023;2(2):e0000198. [FREE Full text] [doi: [10.1371/journal.pdig.0000198](https://doi.org/10.1371/journal.pdig.0000198)] [Medline: [36812645](https://pubmed.ncbi.nlm.nih.gov/36812645/)]
41. Wójcik S, Rulkiewicz A, Pruszczyk P, Lisik W, Poboży M, Domienik-Karłowicz J. Reshaping medical education: performance of ChatGPT on a PES medical examination. *Cardiol J*. 2024;31(3):442-450. [FREE Full text] [doi: [10.5603/cj.97517](https://doi.org/10.5603/cj.97517)] [Medline: [37830257](https://pubmed.ncbi.nlm.nih.gov/37830257/)]
42. Zong H, Li J, Wu E, Wu R, Lu J, Shen B. Performance of ChatGPT on Chinese national medical licensing examinations: a five-year examination evaluation study for physicians, pharmacists and nurses. *BMC Med Educ*. 2024;24(1):143. [FREE Full text] [doi: [10.1186/s12909-024-05125-7](https://doi.org/10.1186/s12909-024-05125-7)] [Medline: [38355517](https://pubmed.ncbi.nlm.nih.gov/38355517/)]
43. Madrid-García A, Rosales-Rosado Z, Freitas-Nuñez D, Pérez-Sancristóbal I, Pato-Cour E, Plasencia-Rodríguez C, et al. Harnessing ChatGPT and GPT-4 for evaluating the rheumatology questions of the Spanish access exam to specialized medical training. *Sci Rep*. 2023;13(1):22129. [FREE Full text] [doi: [10.1038/s41598-023-49483-6](https://doi.org/10.1038/s41598-023-49483-6)] [Medline: [38092821](https://pubmed.ncbi.nlm.nih.gov/38092821/)]
44. Jeyaraman M, Jeyaraman N, Nallakumarasamy A, Yadav S, Bondili SK. ChatGPT in medical education and research: a boon or a bane? *Cureus*. 2023;15(8):e44316. [FREE Full text] [doi: [10.7759/cureus.44316](https://doi.org/10.7759/cureus.44316)] [Medline: [37779749](https://pubmed.ncbi.nlm.nih.gov/37779749/)]
45. Zhang J, Zhang Z. Ethics and governance of trustworthy medical artificial intelligence. *BMC Med Inform Decis Mak*. 2023;23(1):7. [FREE Full text] [doi: [10.1186/s12911-023-02103-9](https://doi.org/10.1186/s12911-023-02103-9)] [Medline: [36639799](https://pubmed.ncbi.nlm.nih.gov/36639799/)]
46. Terms of Use. URL: <https://openai.com/policies/terms-of-use> [accessed 2024-04-30]
47. Liebrez M, Schleifer R, Buadze A, Bhugra D, Smith A. Generating scholarly content with ChatGPT: ethical challenges for medical publishing. *Lancet Digit Health*. 2023;5(3):e105-e106. [FREE Full text] [doi: [10.1016/S2589-7500\(23\)00019-5](https://doi.org/10.1016/S2589-7500(23)00019-5)] [Medline: [36754725](https://pubmed.ncbi.nlm.nih.gov/36754725/)]
48. ChatGPT Generative Pre-trained Transformer, Zhavoronkov A. Rapamycin in the context of pascal's wager: generative pre-trained transformer perspective. *Oncoscience*. 2022;9:82-84. [FREE Full text] [doi: [10.18632/oncoscience.571](https://doi.org/10.18632/oncoscience.571)] [Medline: [36589923](https://pubmed.ncbi.nlm.nih.gov/36589923/)]
49. Marchandot B, Matsushita K, Carmona A, Trimaille A, Morel O. ChatGPT: the next frontier in academic writing for cardiologists or a pandora's box of ethical dilemmas. *Eur Heart J Open*. 2023;3(2):oead007. [FREE Full text] [doi: [10.1093/ehjopen/oead007](https://doi.org/10.1093/ehjopen/oead007)] [Medline: [36915398](https://pubmed.ncbi.nlm.nih.gov/36915398/)]
50. General Data Protection Regulation. URL: <https://gdpr-info.eu/> [accessed 2024-04-18]
51. Personal Information Protection Law. URL: https://www.gov.cn/xinwen/2021-08/20/content_5632486.htm [accessed 2024-04-18]
52. Health Insurance Portability and Accountability Act. URL: <https://www.hhs.gov/hipaa/index.html> [accessed 2024-04-18]
53. Institute of Electrical and Electronics Engineers Global Initiative on Ethics of Autonomous and Intelligent Systems. URL: https://standards.ieee.org/wp-content/uploads/import/documents/other/ead1e_general_principles.pdf [accessed 2024-04-18]
54. European Union's White Paper on Artificial Intelligence. URL: https://www.ouvrirlascience.fr/wp-content/uploads/2020/03/commission-white-paper-artificial-intelligence-feb2020_en.pdf [accessed 2024-04-18]
55. Recommendations on the Ethics of Artificial Intelligence. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000381137> [accessed 2024-04-18]
56. Ethics and Governance of AI in Health: Guidance on Large Multi-modal Models. URL: <https://iris.who.int/bitstream/handle/10665/375579/9789240084759-eng.pdf> [accessed 2024-04-18]
57. Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development. URL: <https://documents.un.org/doc/undoc/ltd/n24/065/92/pdf/n2406592.pdf?token=TozR3GWFrpplwXOHMd&fe=true> [accessed 2024-04-18]
58. El Emam K. Methods for the de-identification of electronic health records for genomic research. *Genome Med*. 2011;3(4):25. [FREE Full text] [doi: [10.1186/gm239](https://doi.org/10.1186/gm239)] [Medline: [21542889](https://pubmed.ncbi.nlm.nih.gov/21542889/)]
59. Amanat A, Rizwan M, Maple C, Zikria YB, Almadhor AS, Kim SW. Blockchain and cloud computing-based secure electronic healthcare records storage and sharing. *Front Public Health*. 2022;10:938707. [FREE Full text] [doi: [10.3389/fpubh.2022.938707](https://doi.org/10.3389/fpubh.2022.938707)] [Medline: [35928494](https://pubmed.ncbi.nlm.nih.gov/35928494/)]
60. de Beijer IAE, van den Oever SR, Charalambous E, Cangioli G, Balaguer J, Bardi E, et al. PanCareSurPass Consortium. IT-related barriers and facilitators to the implementation of a new European eHealth solution, the digital survivorship passport (SurPass version 2.0): semistructured digital survey. *J Med Internet Res*. 2024;26:e49910. [FREE Full text] [doi: [10.2196/49910](https://doi.org/10.2196/49910)] [Medline: [38696248](https://pubmed.ncbi.nlm.nih.gov/38696248/)]
61. Cao Y, Shi Q, Shen L, Chen K, Wang Y, Zeng W, et al. NFTracer: tracing NFT impact dynamics in transaction-flow substitutive systems with visual analytics. *IEEE Trans Vis Comput Graph*. 2024. [doi: [10.1109/TVCG.2024.3402834](https://doi.org/10.1109/TVCG.2024.3402834)] [Medline: [38768002](https://pubmed.ncbi.nlm.nih.gov/38768002/)]

62. King MR, chatGPT. A conversation on artificial intelligence, Chatbots, and plagiarism in higher education. *Cell Mol Bioeng*. 2023;16(1):1-2. [FREE Full text] [doi: [10.1007/s12195-022-00754-8](https://doi.org/10.1007/s12195-022-00754-8)] [Medline: [36660590](https://pubmed.ncbi.nlm.nih.gov/36660590/)]
63. Ong JCL, Seng BJJ, Law JZF, Low LL, Kwa ALH, Giacomini KM, et al. Artificial intelligence, ChatGPT, and other large language models for social determinants of health: current state and future directions. *Cell Rep Med*. 2024;5(1):101356. [FREE Full text] [doi: [10.1016/j.xcrm.2023.101356](https://doi.org/10.1016/j.xcrm.2023.101356)] [Medline: [38232690](https://pubmed.ncbi.nlm.nih.gov/38232690/)]

Abbreviations

AI: artificial intelligence

IP: intellectual property

LLM: large language model

Edited by A Mavragani; submitted 30.04.24; peer-reviewed by MO Khursheed; comments to author 10.06.24; accepted 06.07.24; published 01.08.24

Please cite as:

Zhui L, Fenghe L, Xuehu W, Qining F, Wei R

Ethical Considerations and Fundamental Principles of Large Language Models in Medical Education: Viewpoint

J Med Internet Res 2024;26:e60083

URL: <https://www.jmir.org/2024/1/e60083>

doi: [10.2196/60083](https://doi.org/10.2196/60083)

PMID: [38971715](https://pubmed.ncbi.nlm.nih.gov/38971715/)

©Li Zhui, Li Fenghe, Wang Xuehu, Fu Qining, Ren Wei. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 01.08.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.