



Biomedical Ethical Aspects Towards the Implementation of Artificial Intelligence in Medical Education

Felix Busch^{1,2} · Lisa C. Adams³ · Keno K. Bressem^{1,4}

Accepted: 31 May 2023
© The Author(s) 2023

Abstract

The increasing use of artificial intelligence (AI) in medicine is associated with new ethical challenges and responsibilities. However, special considerations and concerns should be addressed when integrating AI applications into medical education, where healthcare, AI, and education ethics collide. This commentary explores the biomedical ethical responsibilities of medical institutions in incorporating AI applications into medical education by identifying potential concerns and limitations, with the goal of implementing applicable recommendations. The recommendations presented are intended to assist in developing institutional guidelines for the ethical use of AI for medical educators and students.

Keywords Biomedical ethics · AI · Education, Medical · Practice guidelines · Machine learning · ChatGPT

Background

Artificial intelligence (AI) has the potential to revolutionize medicine by improving patient outcomes with personalized medicine, increasing efficiency, and reducing healthcare costs by supporting medical professionals in patient healthcare, diagnosis, decision-making, and conducting research, among many other possible applications [1–5]. Besides, there are several opportunities for how AI applications can be used to enhance medical training. Potential applications include virtual patient simulations or case studies that

allow medical students to practice diagnosing and treating patients in a controlled environment, machine learning for image analysis to help interpret medical images and detect anomalies, natural language processing models for transcribing medical records or content, and tailored or intelligent learning plans and feedback based on personal strengths and weaknesses [6–11]. Moreover, medical students who have received training in AI may be better equipped and more comfortable using AI tools and technologies in their later clinical work as they learn to understand and use AI concepts and their potential applications from the beginning of their careers [12, 13]. On the other hand, large language models (LLMs) such as ChatGPT are becoming increasingly popular and are likely to revolutionize medical education in both positive and negative directions. While LLMs can explain medical terminology to students (and patients) or simulate anamnestic interviews, they could also be used to cheat on exams, papers, and assignments [14]. The increased use of AI in medicine also raises many ethical questions. Common ethical concerns about the use of AI in medicine include lack of transparency, insufficient knowledge about the application used, or false/misleading results [15, 16].

In addition, AI algorithms trained on biased data can lead to incorrect diagnoses and unreasonable or unfair decisions [17]. Hence, AI applications may perpetuate existing inequities in the medical field by providing unequal access to care or making biased treatment decisions. Further concerns include the insufficient protection of data privacy

Lisa C. Adams and Keno K. Bressem contributed equally to this work.

✉ Felix Busch
felix.busch@charite.de

¹ Department of Radiology, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany

² Department of Anesthesiology, Division of Operative Intensive Care Medicine, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany

³ Department of Radiology, Stanford University School of Medicine, Stanford, CA, USA

⁴ Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany

and confidentiality, as well as the lack of informed consent when retrospectively using patient data for training [18, 19]. Finally, there is a risk that AI systems will compromise patient autonomy and dignity by making treatment decisions without appropriate oversight [15, 20].

While current publications on AI-related medical education emphasize key competencies or explore educational AI programs and concepts, they fail to specifically account for the four main pillars of biomedical ethics: autonomy, justice, non-maleficence, and beneficence [21–24]. However, considering biomedical ethics when integrating AI applications into education is of particular importance, especially in the medical field, where vital measures can be taught and healthcare, education, and AI ethics collide. Therefore, this commentary emphasizes ethical issues related to the use of AI applications in the medical curriculum and proposes recommendations for medical institutions within a biomedical ethical framework.

Biomedical Ethical Principles of Using AI in Medical Education

In general, medical institutions should ensure that they have the necessary technical infrastructure, resources, and expertise to support the use of AI in medical education. Therefore, it is required to clearly define the learning purposes and objectives when using educational AI [25]. These include general pedagogical and ethical considerations such as the pedagogical approach and the integrity of teachers and learners, as AI should enhance the classroom experience while preserving the fundamental dimensions of the human being [26]. It may be beneficial to involve all relevant stakeholders, including developers, providers, regulators, faculty, staff, medical bioethicists, and students, in the decision-making process to gather and discuss expertise, desires, concerns, ideas, and moral, legal, and ethical issues and to ensure a satisfactory implementation for all parties within possible limits [27].

This work focuses on the four main principles of biomedical ethics for medical institutions when integrating AI into medical education rather than on the general ethics of AI in education or healthcare.

Autonomy

The principle of autonomy emphasizes the inherent and unconditional value of every human being and their right to self-determination [21, 22]. This includes the ability to make rational judgments and moral choices as well as the unrestricted right to exercise control over one's own decisions. In medical bioethics, autonomy is a cornerstone, guiding the interactions between healthcare providers and patients and

ensuring respect for individual preferences and values. However, integrating AI into medical education poses significant challenges to the autonomy of users, including students, educators, and medical professionals. For example, advanced computational models for natural language processing, such as ChatGPT, occasionally hallucinate, i.e., generate information or responses that are not based on factual data [28]. Although current technological capabilities do not allow the complete elimination of this issue, efforts can be taken to minimize its occurrence. In the context of medical education, it is crucial that LLMs possess mechanisms to transparently signal their limitations or uncertainties to prevent the propagation of erroneous information, ensuring full autonomy (and non-maleficence) among learners. The dependence on AI technology may also result in the lack of development of essential decision-making skills and clinical judgment [29]. Moreover, the complexity and opacity of AI algorithms can make it challenging to comprehend how algorithms arrive at specific decisions, potentially reducing the ability to make informed judgments about the appropriateness and correctness of AI-generated recommendations or diagnoses [30]. Therefore, medical institutions should ensure that the use of AI in the medical curriculum is transparent and comprehensible and that users know when and how different AI applications are used [31]. Furthermore, their informed consent should be obtained to ensure they fully understand and agree to use the application [32].

On the other hand, medical institutions should promote professional responsibility and accountability to empower users to make rational judgments and moral decisions, accounting for the use of AI applications with the best possible knowledge and conscience [33]. Finally, universities should offer AI models as an add-on to the medical curriculum rather than a complete replacement for traditional teaching materials and strategies so that students can decide at any time whether or not to use AI applications.

Justice

The principle of justice comprises fairness, equity, and equal treatment for all individuals [21, 22]. This principle requires that benefits and burdens are distributed equitably among all those affected by a decision or action. Although 60% of the data used to develop AI applications is estimated to be synthetic by 2024, 40% will still be based on real data [34]. This brings forward a key question of justice — how to adequately compensate and acknowledge data sources utilized in training AI models? For medical institutions, this can be challenging to verify, especially for AI products developed by external vendors. Therefore, one potential solution might be to control data-sharing agreements between data providers and vendors before implementing AI applications. On the other hand, new laws and regulations on the use of data in

AI development are needed to set standards and enforce fair practices. Furthermore, before integrating AI into medical education, an equity and social justice framework should be established to avoid a disproportionate impact on certain user populations [35]. This framework should guide the development and application of AI technologies, ensuring that they are accessible to everyone and as individualized as possible to meet the needs and perspectives of users from different backgrounds, social classes, knowledge levels, and interests to ensure that AI is used in a way that is inclusive and respectful of diversity [27, 36]. This can be reached by offering financial assistance, scholarships, or subsidies to ensure that no user is disadvantaged due to lack of access to AI applications or resources, an inclusive and adaptable design of AI applications that is developed on the needs and perspectives of a diverse set of stakeholders, or the use of diverse training data to account for equal treatment of all individuals. For example, to ensure that a broader range of perspectives is considered in the development and implementation of AI technologies, collaborations can be established with educational institutions from diverse socioeconomic backgrounds [37]. In terms of diversity of training data, institutions need to ensure that developers train their algorithms and models on high-quality, diverse data that accurately represent the population being studied so that the resulting educational content is unbiased and fair [38, 39]. This requires careful attention to data collection and curation, as well as ongoing monitoring and refinement of algorithms. Finally, AI tools that are used in medical education should undergo regular audits to detect and correct potential biases and inequities that may inadvertently arise, which also addresses non-maleficence, for example, when exploring treatment recommendations for patients [40].

Non-maleficence

Non-maleficence is the principle that emphasizes the importance of not causing harm and minimizing potential negative consequences [21, 22]. Using AI in medicine presents both opportunities and challenges to uphold the principle of non-maleficence. Especially AI algorithms in medical education must be carefully designed, validated, and evaluated to ensure that they produce accurate and reliable results that do not mislead users and consecutively put patients at risk. Importantly, all users should be adequately educated about the application's functionalities, known biases, and potential risks prior to using AI models, which is an essential prerequisite for all four bioethical principles [41]. For example, when AI algorithms are trained based on biased or incomplete data, they can perpetuate or reinforce existing biases and inequities in healthcare. This may lead to inaccurate or discriminatory outcomes and potential harm to specific patient groups, such as the recently discovered bias against

women in AI tools [42]. To adequately assess any limitations or biases of medical AI applications, complete and detailed information and understanding of the training data for each application are required, as described above. However, this is often not applicable, for example, when applications are not open source or with missing knowledge about the technology used [43]. Therefore, medical institutions should outline the limitations of any AI application used and point out the risks of applications without transparency and of systems trained on biased or unrepresentative datasets. In this regard, medical data experts, such as physicians and medical educators, have a special ethical responsibility, as they are the ones who can detect bias in data, validate models, and train students in AI sufficiently to detect errors themselves.

On the other hand, medical institutions should provide AI-independent information to enable students to make informed decisions without the use of AI or to compare AI-generated data (which also reinforces autonomy), particularly when the results of AI applications appear to be divergent or misleading, for example, when researching diagnoses or treatment recommendations [44]. Moreover, students must be aware that AI-generated recommendations may not provide a complete and all-encompassing understanding of a patient's unique medical situation. It also should be noted that the different AI-generated information also poses different risks when the retrieved information is applied to patients. For example, while simply retrieving information about drug interactions or dosing regimens using LLMs poses a lower risk, extended use of the models to the point of identifying symptoms, diagnosis, and treatment planning without AI-independent verification can be particularly dangerous [45].

When AI applications are used to work with real patient data, privacy and confidentiality of sensitive data must be warranted, and informed consent might be required from patients whose data are being used [46]. This requires ensuring that all users adhere to strict privacy regulations and policies, for example, the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA) for institutions in the USA. In addition, approval from the Food and Drug Administration (FDA), European Medicines Agency (EMA), or country-specific equivalents should be obtained for each new use case, as well as approval from the internal institutional review board, including an appropriate approach to ensure privacy and confidentiality. The use of AI applications for individual medical education or knowledge acquisition will also generate new sensitive data, for instance, the assessment and analysis of individual performance, which may lead to competitive pressures or negative emotions when results are compared with those of fellow students or when faculty members are provided with access [47, 48]. Therefore, institutions should handle individual results with confidentiality.

Principles of biomedical ethics			
Autonomy	Justice	Non-maleficence	Beneficence
Ensure that AI applications are transparent and comprehensible.	Ensure equal access to AI applications and resources.	Emphasize the risks of bias and discrimination and encourage critical thinking in the use of AI.	Provide comprehensive training before using AI applications.
Obtain informed consent from all users.	Ensure that AI applications are tailored to diverse user needs and backgrounds.	Provide AI-independent resources for verification.	Work together in close collaboration with all stakeholders.
Promote professional responsibility and accountability.	Ensure the use of diverse training data.	Secure the privacy and confidentiality of sensitive data.	Provide easy access to evaluation.
Implement AI models as an add-on to the medical curriculum.	Implement regular audits of the AI tools in use.	Regularly monitor potential risks and harms associated with AI applications or new AI use cases.	Systematically observe and reevaluate the impact of AI applications to optimize their use in medical education.

Fig. 1 Overview of key recommendations for medical institutions towards integrating artificial intelligence (AI) into medical education, based on the four core principles of biomedical ethics — autonomy, justice, non-maleficence, and beneficence

This may include decentralized storage of data and evaluation of individual skills, which can only be accessed by the respective individual.

Finally, AI in medicine is a rapidly evolving area of research. This can certainly lead to dissatisfaction and be disruptive for medical institutions. However, constantly changing applications or expanded use cases may also lead to newly identified risks or harms associated with AI, such as privacy or security concerns, discriminatory outcomes, and embedded or inserted biases [49]. Therefore, ethical guidelines should be continuously monitored and adjusted.

Beneficence

The principle of beneficence in biomedical ethics emphasizes the obligation to promote and protect the welfare and well-being of individuals [21, 22]. In the context of AI in medical education, beneficence involves providing appropriate training on AI applications before their implementation to maximize the benefits for all stakeholders, including students, educators, and experts [7, 50]. Moreover, as outlined above, appropriate education and training about AI algorithms not only positively influence beneficence but also empower autonomy, justice, and non-maleficence by enabling more informed decision-making, consideration of inequalities and biases, and more effective integration of AI into medical education. The training on AI may encompass tutorials, supervised workshops, and offline or online materials that help users understand how and when to use each application. On the other hand, working

together in close collaboration with all stakeholders, including students, educators, AI developers, and health-care professionals, is essential for the most beneficial implementation of AI in medical education, accounting for special needs among different disciplines, for example, when exploring which AI application can enhance the learning experience for each specialty [29]. Finally, the regular assessment and optimization of AI applications in medical education are crucial to ensure that these technologies are employed in a way that maximizes benefits (and minimizes potential harms) [51]. By systematically monitoring and reevaluating the impact of AI applications on the medical curriculum, for example, by providing an easily accessible tool for evaluation, institutions can make data-driven decisions on how to improve or modify the use of AI to better serve the educational needs of students and educators.

Conclusions

A summary of all the recommendations developed within this comment can be viewed in Fig. 1. While integrating AI in medical education has the potential to provide a more immersive, interactive, and personalized learning experience, it is essential to establish a biomedical ethical framework beforehand, which should be regularly reevaluated and optimized based on user feedback and the latest developments in the field.

Acknowledgements Keno K. Bressem is grateful for his participation in the Berlin Institute of Health (BIH) Charité Digital Clinician Scientist Program, funded by the Charité – Universitätsmedizin Berlin and the BIH.

Author Contribution Conceptualization: Felix Busch, Lisa C. Adams, Keno K. Bressem. Project administration: Felix Busch; Resources: Felix Busch, Lisa C. Adams, Keno K. Bressem. Writing — original draft: Felix Busch. Writing — review and editing: Felix Busch, Lisa C. Adams, Keno K. Bressem.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Ethics Approval Not applicable.

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Curr Cardiol Rep*. 2014;16(1):441. <https://doi.org/10.1007/s11886-013-0441-8>.
- Mertz L. AI tools poised to improve patient health care. *IEEE Pulse*. 2022;13(2):2–6. <https://doi.org/10.1109/mpuls.2022.3159038>.
- Amato F, López A, Peña-Méndez EM, Vañhara P, Hampl A, Havel J. Artificial neural networks in medical diagnosis. *J Appl Biomed*. 2013;11(2):47–58. <https://doi.org/10.2478/v10136-012-0031-x>.
- Bennett CC, Hauser K. Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach. *Artif Intell Med*. 2013;57(1):9–19. <https://doi.org/10.1016/j.artmed.2012.12.003>.
- Mosch L, Fürstenau D, Brandt J, et al. The medical profession transformed by artificial intelligence: qualitative study. *Digit Health*. 2022;8:20552076221143904. <https://doi.org/10.1177/20552076221143903>.
- Wood EA, Ange BL, Miller DD. Are we ready to integrate artificial intelligence literacy into medical school curriculum: students and faculty survey. *J Med Educ Curric*. 2021;8:23821205211024080. <https://doi.org/10.1177/23821205211024078>.
- Chan KS, Zary N. Applications and challenges of implementing artificial intelligence in medical education: integrative review. *JMIR Med Educ*. 2019;5(1):e13930. <https://doi.org/10.2196/13930>.
- Lillehaug S-I, Lajoie SP. AI in medical education—another grand challenge for medical informatics. *Artif Intell Med*. 1998;12(3):197–225. [https://doi.org/10.1016/S0933-3657\(97\)00054-7](https://doi.org/10.1016/S0933-3657(97)00054-7).
- Li YS, Lam CSN, See C. Using a machine learning architecture to Create an AI-powered chatbot for anatomy education. *Med Sci Educ*. 2021;31:1729–30. <https://doi.org/10.1007/s40670-021-01405->.
- Nagy M, Radakovich N, Nazha A. Why machine learning should be taught in medical schools. *Med Sci Educ*. 2022;32:529–32. <https://doi.org/10.1007/s40670-022-01502-3a>.
- Adams LC, Truhn D, Busch F, Kader A, et al. Leveraging GPT-4 for post hoc transformation of free-text radiology reports into structured reporting: a multilingual feasibility study. *Radiology*. 2023. <https://doi.org/10.1148/radiol.230725>.
- Sit C, Srinivasan R, Amlani A, et al. Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: a multicentre survey. *Insights Imaging*. 2020;11(1):14. <https://doi.org/10.1186/s13244-019-0830-7>.
- Masters K. Artificial intelligence in medical education. *Med Teach*. 2019;41(9):976–80. <https://doi.org/10.1080/0142159X.2019.1595557>.
- Kung TH, Cheatham M, Medinilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *medRxiv*. 2022. <https://doi.org/10.1101/2022.12.19.22283643>.
- Rigby MJ. Ethical dimensions of using artificial intelligence in health care. *AMA J Ethics*. 2019;21(2):121–4. <https://doi.org/10.1001/amajethics.2019.121>.
- Sullivan HR, Schweikart SJ. Are current tort liability doctrines adequate for addressing injury caused by AI? *AMA J Ethics*. 2019;21(2):160–6.
- Chen IY, Szolovits P, Ghassemi M. Can AI help reduce disparities in general medical and mental health care? *AMA J Ethics*. 2019;21(2):E167–179. <https://doi.org/10.1001/amajethics.2019.167>.
- Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. *Artif Intell Med*. 2014;62(1):1–10. <https://doi.org/10.1016/j.artmed.2014.06.004>.
- Kargl M, Plass M, Müller H. A literature review on ethics for ai in biomedical research and biobanking. *Yearb Med Inform*. 2022;31(1):152–60. <https://doi.org/10.1055/s-0042-1742516>.
- Wilhelm D, Hartwig R, McLennan S, et al. Ethical, legal and social implications in the use of artificial intelligence-based technologies in surgery: principles, implementation and importance for the user. *Chirurg*. 2022;93(3):223–33. <https://doi.org/10.1007/s00104-022-01574-2>.
- Beauchamp TL, Childress JF. Principles of biomedical ethics. 8th ed. Oxford: Oxford Publishing Press; 2019. p. 1–512.
- Varkey B. Principles of clinical ethics and their application to practice. *Med Princ Pract*. 2021;30:17–28. <https://doi.org/10.1159/000509119>.
- Çalışkan SA, Demir K, Karaca O. Artificial intelligence in medical education curriculum: an e-Delphi study for competencies. *PLoS ONE*. 2022;17(7):e0271872. <https://doi.org/10.1371/journal.pone.0271872>.
- Charow R, Jeyakumar T, Younus S, et al. Artificial intelligence education programs for health care professionals: scoping review. *JMIR Med Educ*. 2021;7(4):e31043. <https://doi.org/10.2196/31043>.
- Holmes W, Porayska-Pomsta K, Holstein K, et al. Ethics of AI in education: towards a community-wide framework. *Int J Artif Intell Educ*. 2022;32(3):504–26. <https://doi.org/10.1007/s40593-021-00239-1>.

26. Aiken R, Epstein R. Ethical guidelines for AI in education: starting a conversation. *Int J Artif Intell Educ*. 2000;11:163–76.
27. Schiff D. Education for AI, not AI for education: the role of education and ethics in national AI policy strategies. *Int J Artif Intell Educ*. 2022;32(3):527–63. <https://doi.org/10.1007/s40593-021-00270-2>.
28. Dziri N, Milton S, Yu M, Zaiane O, Reddy S. On the origin of hallucinations in conversational models: Is it the datasets or the models? *arXiv*. 2022;arXiv:220407931. <https://doi.org/10.48550/arXiv.2204.07931>.
29. Amann J, Blasimme A, Vayena E, et al. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak*. 2020;20:310. <https://doi.org/10.1186/s12911-020-01332-6>.
30. Durán JM, Jongsma KR. Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. *J Med Ethics*. 2021;47:329–35. <https://doi.org/10.1136/medethics-2020-106820>.
31. Shin DD. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *Int J Hum Comput Stud*. 2021;146:102551. <https://doi.org/10.1016/j.ijhcs.2020.102551>.
32. Bogina V, Hartman A, Kuflik T, Shulner-Tal A. Educating software and ai stakeholders about algorithmic fairness, accountability, transparency and ethics. *Int J Artif Intell Educ*. 2022;32(3):808–33. <https://doi.org/10.1007/s40593-021-00248-0>.
33. Kassim PNJ, Osman A, Muhammad RW. Educating future medical professionals with the fundamentals of law and ethics. *Int Med J Malays*. 2020;16(2). <https://doi.org/10.31436/IMJM.V16I2.334>.
34. Castellanos S. Fake it to make it: companies beef up AI models with synthetic data. *The Wallstreet Journal*. New York, NY: Dow Jones & Company, Inc., 2021. <https://www.wsj.com/articles/fake-it-to-make-it-companies-beef-up-ai-models-with-synthetic-data-11627032601>. Accessed 15 May 2023.
35. Coria AL, McKelvey T, Charlton P, et al. The design of a medical school social justice curriculum. *Acad Med*. 2013;88:1442–9. <https://doi.org/10.1097/ACM.0b013e3182a325be>.
36. Dennis M, Masthoff J, Mellish C. Adapting progress feedback and emotional support to learner personality. *Int J Artif Intell Educ*. 2016;26(3):877–931. <https://doi.org/10.1007/s40593-015-0059-7>.
37. Hagerty A, Rubinov I. Global AI ethics: a review of the social impacts and ethical implications of artificial intelligence. *arXiv*. 2019;arXiv:1907.07892. <https://doi.org/10.48550/arXiv.1907.07892>.
38. Esaki T, Watanabe R, Kawashima H, et al. Data curation can improve the prediction accuracy of metabolic intrinsic clearance. *Mol Inform*. 2019;38(1–2):1800086. <https://doi.org/10.1002/minf.201800086>.
39. Bozkurt S, Cahan EM, Seneviratne MG, et al. Reporting of demographic data and representativeness in machine learning models using electronic health records. *J Am Med Inform Assoc*. 2020;27(12):1878–84. <https://doi.org/10.1093/jamia/ocaa164>.
40. Esmaeilzadeh P, Mirzaei T, Dharanikota S. Patients' perceptions toward human-artificial intelligence interaction in health care: experimental study. *J Med Internet Res*. 2021;23(11):e25856. <https://doi.org/10.2196/25856>.
41. Solomonides AE, Koski E, Atabaki SM, et al. Defining AMIA's artificial intelligence principles. *J Am Med Inform Assoc*. 2022;29(4):585–91. <https://doi.org/10.1093/jamia/ocac006>.
42. Dastin J. Amazon scraps secret AI recruiting tool that showed bias against women. In: Martin K, editor. *Ethics of Data and Analytics*. 1st ed. Boca Raton, FL: CRC Press; 2018. p. 296–9.
43. Bellamy RKE, Dey K, Hind M, et al. AI Fairness 360: an extensible toolkit for detecting and mitigating algorithmic bias. *IBM J Res Dev*. 2019;63(4/5):4:1–4:15. <https://doi.org/10.1147/JRD.2019.2942287>.
44. Ryan M. In AI we trust: ethics, artificial intelligence, and reliability. *Sci Eng Ethics*. 2020;26(5):2749–67. <https://doi.org/10.1007/s11948-020-00228-y>.
45. WEF. Chatbots RESET. A framework for governing responsible use of conversational AI in healthcare. http://www3.weforum.org/docs/WEF_Governance_of_Chatbots_in_Healthcare_2020.pdfReturn. Accessed 10 Dec 2022.
46. Murdoch B. Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC Med Ethics*. 2021;22(1):122. <https://doi.org/10.1186/s12910-021-00687-3>.
47. Graesser AC, D'Mello SK, Strain AC. Emotions in advanced learning technologies. In: Linnenbrink-Garcia L, Pekrun R, editors. *International handbook of emotions in education*. Abingdon-Thames: Routledge; 2014. p. 473–93.
48. Anwar M, Greer J. Facilitating trust in privacy-preserving E-learning environments. *TLT*. 2012;5(1):62–73. <https://doi.org/10.1109/TLT.2011.23>.
49. Zhou J, Chen F, Berry A, Reed M, Zhang S, Savage S, editors. A survey on ethical principles of AI and implementations. *IEEE SSCI*. 2020;1–4. <https://doi.org/10.1109/SSCI47803.2020.9308437>.
50. Sapci AH, Sapci HA. Artificial intelligence education and tools for medical and health informatics students: systematic review. *JMIR Med Educ*. 2020;6(1):e19285. <https://doi.org/10.2196/19285>.
51. Rogers WA, Draper H, Carter SM. Evaluation of artificial intelligence clinical applications: detailed case analyses show value of healthcare ethics approach in identifying patient care issues. *Bioethics*. 2021;35(7):623–33. <https://doi.org/10.1111/bioe.12885>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.