

Risks of Artificial Intelligence (AI) in Medicine

Nikolaos Siafakas^{1*}, Eirini Vasarmidi^{2*}

Artificial Intelligence (AI) is the most rapidly advancing science affecting the majority of human activities in a rather beneficial way. AI has modified human life improving communications, traveling, car industry, commerce, construction, and agriculture^{1,2}. AI using Deep Reinforcement Learning Techniques is rapidly progressing and it is predicted that by the year of 2050, or even earlier, it will reach the level of general human intelligence³. Although all this progress is welcomed, it is foreseen that it may be followed by significant threats to humanity^{4,5}.

Medicine is one of the sciences that have been beneficially affected by improved accuracy of the diagnosis, epidemiology, staging-severity, prognosis, and treatment of numerous diseases^{6,7}. In addition, AI has played a crucial role in the discovery of new drugs and revolutionized medical education, medical information, and medical practice. It is obvious that the aim of AI in medicine is to promote human welfare, equality and advanced healthcare and preserving the autonomy security and transparency of the data of individual patients. However, the development and implementation of AIs present significant risks in the field of practicing medicine as well as in biological research^{8,9}.

The following risks and dangers will be discussed: 1) as exist at present, misuse of the data, risks during the development and/or implementation of AI, risks of medical education; 2) as may present shortly, replacement of various medical tests and examinations or medical specialties; and 3) when AI may become stronger (Super AI) than human intelligence. Finally, a proposal based on Medical Ethics will be presented for AI to remain beneficial to humanity in the future¹⁰.

Risks from AI in medicine: present

Data

AI systems use specific algorithms that need large datasets to improve their accuracy (specificity/sensitivity). This process is at great risk, as far as the security, privacy, and confidentiality of the sensitive individual patient's data, is concerned¹¹. Today, the danger of hacking of such datasets has increased tremendously due to the interest of the pharmaceutical or insurance companies. In addition, the hacking of medical files could be a part of a cyber-attack against a government^{12,13}.

Another issue is data bias. During the collection of the data, intentionally or unintentionally, certain minorities, races, ethnicities, or genders may be significantly misrepresented. Therefore, these algorithms are biased and inadequately represent the general population^{14,15}. This bias effect could be magnified by the reluctance of medical practitioners, hospitals, or other health organizations, to provide the medical files of their patients due to fears of security leaks. Another significant danger of medical data misuse is the data poisoning effect, which refers to the deliberate manipulation of medical data to introduce errors or biases in healthcare. This has serious consequences on the accuracy and reliability of medical recommendations. This could also affect the outcomes of clinical trials or insurance claims¹¹. Finally, when AI uses different epidemiological data models, as was seen during the COVID-19 epidemic, this could lead to different conclusions.

Development of AI algorithms in medicine

An inaccurate medical algorithm could affect a large number of patients. This

AFFILIATION

1 Department of Computer Science, University of Crete, Heraklion, Greece

2 Department of Respiratory Medicine and Laboratory of Molecular and Cellular Pneumology, Medical School, University of Crete, Heraklion, Greece

* Contributed equally

CORRESPONDENCE TO

Nikolaos Siafakas. Department of Computer Science, Room H307, Computer Science Building, University Campus, Heraklion, 70013, Crete, Greece.

E-mail: siafakan@uoc.gr

ORCID iD: <https://orcid.org/0000-0001-6747-2377>

KEYWORDS

digital mind, medical education, dangers, super-AI, ethics, health sciences

Received: 19 June 2024

Revised: 22 July 2024

Accepted: 28 July 2024

may be the result of inadequate testing of such an algorithm since there are currently no solid rules to test its validity. A potential example is the case of double-blind trials, which are generally considered the most powerful tool to establish the effect of a particular treatment¹⁶. Nevertheless, it is unclear who would be responsible in case of such a mistake: the personal doctor, the hospital, the provider of the machine, or the company developing the algorithm. Therefore, there are significant legal issues of accountability for medical errors due to malfunction of the machines¹⁵.

Implementation

Health providers are not adequately trained to use AI technologies, thus, the implementation of these modes of treatment in practice, raises additional security and error issues, until a widely accepted training system for healthcare providers can be established¹⁷. Furthermore, AI may harm the doctor–patient relationship¹⁸. In-depth knowledge of AI evaluation and performance is required by doctors, to communicate the role of AI meaningfully to patients, and avoid patients' fear and confusion¹⁹.

Fake news-medical information

Algorithms of fake information have been used to affect human behavior altering the socio-economic patterns of individuals or societies. In that case, AI generates articles and social media posts that are designed to deceive the public since they appear legitimate but are completely fabricated. It was reported that fake news could affect even the results of elections^{20,21}. Similarly, AI has been used to generate, amplify, and manipulate the perception of medical issues spreading confusion. An example is the use of fake news in the global movement against vaccinations. People without the appropriate background of knowledge and training can easily misunderstand and misinterpret information about medicine and science. Unfortunately, fake information may affect even health practitioners²².

Risks in medical education

Although AI made it extremely easy to access medical information, the crucial concern is that this information is not always properly validated. The recently developed Chatboxes, which can provide written articles, answers to university exams, or the contents of a requested lecture, should be also considered as potential threats of AI¹⁹. As their accuracy and validity in various medical aspects have not yet been vigilantly tested, they may represent significant threats.

Moreover, there is the so-called 'lazy doctor' effect^{18,23}. If the doctor exclusively uses AI algorithms, for the diagnosis as well as the treatment options, this may lead progressively to catastrophic forgetting of practical skills and loss of intellectual creativity and capacity to solve medical problems, taking into consideration all the medical and ethics parameters of the patients.

Risks of AI in medicine: near future

During the development phase (approximately up to 2050), in terms of social gaps, it is expected that AI in medicine, instead of promoting equality in health services around the globe, may increase the gap between technologically advanced countries and those lacking digital facilities.

Furthermore, the race of the major AI companies to develop the best, fastest, cheapest, and most profitable medical algorithms, raises extreme pressure on the developing scientists and that may cause fatal errors and a lack of transparency^{14,24}. Therefore, even AI is still weak during this period, all the previously discussed risks are expected to be intensified. Deeper applications of AI in biology and genetics may lead to dangerous pathways for humanity since AI is going to manipulate human biological functions and genetic codes²⁵⁻²⁷. Currently, scientists are trying to connect machines with human brains or implant digital chips into humans. This field is known as brain-computer interface (BCI) technology, which aims to create a direct communication pathway between the brain and external devices, such as computers or prosthetic body parts, primarily used for medical purposes^{28,29}. Although this technology seems fascinating in assisting individuals with disabilities, it is important to note the ethical considerations that are raised^{30,31}.

In addition, as it is predicted that AI will cause huge global unemployment, this would affect also the healthcare professions²⁴. The warning about the replacement of tests and examinations, as well as particular medical specialties by AI, is evident. It has been claimed that specialties based on pattern recognition such as Radiology, Pathology, Cytology, Microbiology, and Dermatology could be the first to be substituted by AI^{25,32}. AI algorithms have already been studied in several medical fields to provide an easier way for diagnosis and evaluation of treatment response³².

Risks of AI in medicine: distant future

Currently, AI is weaker than human intelligence, but as AI is very rapidly advancing, it is foreseen that it will reach the general level of human intelligence by the year 2050 or even earlier^{33,34}. Thereafter, AI easily would become Super AI, an AI stronger than the human brain. Super AI would have general intelligence and might possess the capability to improve its algorithms, resulting in rapid advances. When (and if) this happens, even if this could be the most significant scientific event, it raises unpredictable issues of how the machines will use medical science³⁴. If, finally, AI algorithms replace doctors' evaluations, there is a rationale to wonder whether all medical doctors are going to be replaced by robots using AI algorithms. Although this is an extremely difficult question to be answered at present, some experts argue that this may happen in the distant future.

It is even more difficult to predict accurately how Super AI will help scientists interfere with the human genome or use new drugs on humans. Genetic editing may be done

using biotechnological methods, while manipulation of the human brain may be achieved via various means, including psychological manipulation, sophisticated brain-computer interfaces, or even direct neural interference^{29,35,36}. Even if it sounds like scientific fiction, AI may produce clones of 'human laborers' by altering human genetics. There are scenarios describing how machines either intentionally or unintentionally may use medicine to harm humans. In the future, Super AI may try to explore the human brain to advance its scientific knowledge. Super AI may behave towards humans the way we have behaved to animals.

Finally, there is a large number of unknown risks and dangers or never thought of scenarios of how Super AI will use medicine in the distant future^{1,25,27}. To avoid the possible harmful effects of AI that may lead to human extinction, it is of paramount importance to develop a robust control system for safe and human-friendly Super AI^{10,37}.

Ethical pathway for safe AI

There are several proposals to maintain AI-friendly and beneficial for the human race, in the future. Our proposal is based on the idea that medical ethics could be the prototype for digital ethics. Using deep machine learning techniques, we can embed the machines with ethical and moral principles¹⁰.

Medical Ethics was the first attempt at human morals started 2500 years ago with the Hippocratic Oath. In modern times the Oath is complimented by internationally accepted ethics rules and the various ethics committees. Analogously, AI ethics could have similar components: A Hippocratic oath for computer scientists³⁸, globally accepted ethics rules, and ethics committees for digital science^{10,37}.

The Oath is expected to improve the ethics knowledge in the computer society and particularly of the AI scientists, who will be responsible for enriching the machines with ethics algorithms. In addition, the Oath may emphasize the personal responsibilities of each scientist, acquainting them with the consequences of wrongdoing³⁸. Details of this proposal have been previously presented in detail^{10,38}. Internationally accepted ethics and moral rules are urgently needed and if under the auspices, for example, of the United Nations it may be easier to implement globally. Complimentary Ethics Committees for AI at Universities, Research Institutes, and AI companies may add moral filters towards safe AI, enhancing the responsibility and accountability of scientists.

The history of science shows that major discoveries are followed by greater risks. Nuclear power, climatic changes, and AI are considered human inventions that may lead to the extinction of humanity^{1,3,4}. In this article, the currently existing risks and dangers of AI in medicine, as well as the potential ones in the near and distant future, were discussed. Currently, the major risk concerns the safety and confidentiality of the sensitive medical data of each patient. Even though computer scientists are making tremendous

efforts to develop secure methods, medical organizations are reluctant to provide such data, while medical doctors are also not fully convinced. This looks like the fight of policing against crime that is not always successful, although computer scientists are trying vigorously to develop strategies to detect and prevent data manipulations and advance robust cyber security and validation techniques.

Data bias, in other words, incomplete data, or intentional poisoning of the data, are other known dangers of the use of AI in medicine. Incomplete data sets may become harmful when applied to various minorities, such as ethnic, gender, or race that have been excluded from the data. Poisoning of medical data could be made for obvious insurance-claiming reasons, or for affecting the results of clinical trials³¹. Moreover, possible errors of AI in medicine have extremely complicated legal consequences concerning responsibility. Due to the lack of transparency in the development of medical algorithms and the difficulties in the implementation by healthcare professionals, it would be very difficult to verify accountability and apply for claims or penalties^{31,37,39}. International organizations and governments are now trying to fill these gaps in legislation, with the European Union presenting recently the most comprehensive document^{40,41}. The inform consent forms, when needed, pose another significant threat because it is difficult to explain complicated AI procedures and test understandability.

Additionally, it is predicted that AI is going to affect drastically medical education and medical practice. Therefore, Medical Schools and Medical Societies should closely monitor these changes and modify their curriculum accordingly. A possible replacement of various medical specialties by AI may lead to unemployment among health practitioners. Another alarming danger is that in the future Medical Doctors may become over-reliant on AI. This could lead to a decrease in critical thinking, clinical judgment, and their ability to solve problems. Thus, the syndrome of 'lazy student' and 'lazy doctor' may increase^{18,23}.

The future direction is Super AI that will be able to perform any intellectual task. This remains a theoretical concept for now, but the possibility of Super AI development causes debates and concerns among scientists and experts in AI and ethics, given the idea that in such a case, Super AI may manipulate the human brain. Importantly, editing of the human genome by AI could have catastrophic consequences of unpredicted dimensions for humanity. When (and if) machines reach the level of Super AI, it will be extremely difficult to predict how they will use their medical and genetic scientific knowledge on humans^{1,4,25}.

While AI is at its developmental stage, it is of utmost importance to act in a preventive way and develop solid and safe mechanisms to ensure that AI remains human-friendly in the future^{1,37}. Public awareness is crucial, but apart from that, robustly safe mechanisms should be developed by scientists for AI to remain beneficial to humanity, well before it becomes stronger than humans. One way to achieve a

friendly AI is to embed the machines with ethics algorithms. We argue that if Medical Ethics become the prototype for Digital Morals, then the dangers of AI on general human life, and on medicine in particular, could be minimized. It was proposed that an oath similar to the Hippocratic one could be the foundation for Computer Ethics³⁸. Such an oath will enhance the ethical and moral responsibilities of AI scientists and their knowledge about the consequences of wrong practicing. Finally, this oath for Computer scientists will increase global awareness of the potential misuse of AI in medicine and will enhance international collaboration for a safe AI^{10,38}. As the advance of science leads to the ongoing enormous development of AI, the key for humanity is to accomplish symbiosis with AI.

Conclusion

Although AI has a huge beneficial impact on medical science, it is followed by several significant risks and dangers. It is strongly suggested for medical organizations to monitor the changes which are associated with the giant steps of AI development, and modify accordingly medical education and practice. The major risks might emerge when AI becomes more powerful than the human brain, thus it is of paramount importance to develop solid and safe mechanisms to keep AI under control. The establishment of an ethical pathway could be one of the safe ways for AI to remain human-friendly in the future.

CONFLICTS OF INTEREST

The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. The authors declare that they have no competing interests, financial or otherwise, related to the current work. E. Vasarmidi reports that in the past 36 months, she received the ERS Long Term Research Fellowship and the HTS Research Grant, which are not related to this manuscript. Additionally, she received support for attending meetings from Menarini, Chiesi, and BI.

FUNDING

There was no source of funding for this research.

ETHICAL APPROVAL AND INFORMED CONSENT

Ethical approval and informed consent were not required for this study.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created.

PROVENANCE AND PEER REVIEW

Not commissioned; externally peer reviewed.

REFERENCES

1. Bostrom N. Ethical Issues in Advanced Artificial Intelligence.

In: Wallach W, Asaro P, eds. Machine Ethics and Robot Ethics. Routledge; 2017.

2. Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *N Engl J Med*. 2023;388(13):1233-1239. doi:[10.1056/NEJMs2214184](https://doi.org/10.1056/NEJMs2214184)

3. Harari YN. *Homo Deus: A Brief History of Tomorrow*. Harvill Secker; 2017.

4. Kurzweil R. *How to Create a Mind: The Secret of Human Thought Revealed*. Prelude; 2014.

5. McLean S, Read GJM, Thompson J, Baber C, Stanton NA, Salmon PM. The risks associated with Artificial General Intelligence: a systematic review. *J Exp Theor Artif Intell*. 2023;35(5):649-663. doi:[10.1080/0952813X.2021.1964003](https://doi.org/10.1080/0952813X.2021.1964003)

6. Thomas LB, Mastorides SM, Viswanadhan NA, Jakey CE, Borkowski AA. Artificial Intelligence: review of current and future applications in medicine. *Fed Pract*. 2021;38(11):527-538. doi:[10.12788/fp.0174](https://doi.org/10.12788/fp.0174)

7. Pagallo U, O'Sullivan S, Nevejans N, et al. The underuse of AI in the health sector: opportunity costs, success stories, risks and recommendations. *Health Technol (Berl)*. 2024;14(1):1-14. doi:[10.1007/s12553-023-00806-7](https://doi.org/10.1007/s12553-023-00806-7)

8. Zaidi SF, Shaikh A, Surani S. The pulse of AI: implementation of Artificial Intelligence in healthcare and its potential hazards. *Open Respir Med J*. 2024;18:e18743064289936. doi:[10.2174/0118743064289936240115105057](https://doi.org/10.2174/0118743064289936240115105057)

9. Kiener M. Artificial intelligence in medicine and the disclosure of risks. *AI Soc*. 2021;36(3):705-713. doi:[10.1007/s00146-020-01085-w](https://doi.org/10.1007/s00146-020-01085-w)

10. Sifakakos NM. Medical ethics prototype for Artificial Intelligence ethics. *Journal of Philosophy and Ethics*. 2023;5(2):1-3. doi:[10.22259/2642-8415.0502001](https://doi.org/10.22259/2642-8415.0502001)

11. Alonso A, Siracuse JJ. Protecting patient safety and privacy in the era of artificial intelligence. *Semin Vasc Surg*. 2023;36(3):426-429. doi:[10.1053/j.semvasc Surg.2023.06.002](https://doi.org/10.1053/j.semvasc Surg.2023.06.002)

12. Dunn Cavelt M. Breaking the cyber-security dilemma: aligning security needs and removing vulnerabilities. *Sci Eng Ethics*. 2014;20(3):701-715. doi:[10.1007/s11948-014-9551-y](https://doi.org/10.1007/s11948-014-9551-y)

13. Herrin J, Dempsey BJ III. Web-enabled medical databases: a threat to security? *Methods Inf Med*. 2000;39(04/05):298-302. doi:[10.1055/s-0038-1634448](https://doi.org/10.1055/s-0038-1634448)

14. Rodgers CM, Ellingson SR, Chatterjee P. Open data and transparency in artificial intelligence and machine learning: a new era of research. *F1000Res*. 2023;12:387. doi:[10.12688/f1000research.133019.1](https://doi.org/10.12688/f1000research.133019.1)

15. Kiseleva A, Kotzinos D, De Hert P. Transparency of AI in healthcare as a multilayered system of accountabilities: between legal requirements and technical limitations. *Front Artif Intell*. 2022;5:879603. doi:[10.3389/frai.2022.879603](https://doi.org/10.3389/frai.2022.879603)

16. Park HJ. Patient perspectives on informed consent for medical AI: a web-based experiment. *Digit Health*. 2024;10:1-16. doi:[10.1177/20552076241247938](https://doi.org/10.1177/20552076241247938)

17. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing Artificial Intelligence training in

- medical education. *JMIR Med Educ.* 2019;5(2):e16048. doi:[10.2196/16048](https://doi.org/10.2196/16048)
18. Sauerbrei A, Kerasidou A, Lucivero F, Hallowell N. The impact of artificial intelligence on the person-centred, doctor-patient relationship: some problems and solutions. *BMC Med Inform Decis Mak.* 2023;23(1):73. doi:[10.1186/s12911-023-02162-y](https://doi.org/10.1186/s12911-023-02162-y)
 19. Banja JD, Hollstein RD, Bruno MA. When Artificial Intelligence models surpass physician performance: medical malpractice liability in an era of advanced Artificial Intelligence. *J Am Coll Radiol.* 2022;19(7):816-820. doi:[10.1016/j.jacr.2021.11.014](https://doi.org/10.1016/j.jacr.2021.11.014)
 20. How worried should you be about AI disrupting elections? *The Economist.* August 31, 2023. Accessed June 15, 2024. <https://www.economist.com/leaders/2023/08/31/how-artificial-intelligence-will-affect-the-elections-of-2024>
 21. How election experts are thinking about AI and its impact on the 2024 elections. Johns Hopkins University. June 5, 2024. Accessed June 15, 2024. <https://hub.jhu.edu/2024/06/05/election-experts-artificial-intelligence/>
 22. Park HJ. The rise of generative artificial intelligence and the threat of fake news and disinformation online: perspectives from sexual medicine. *Investig Clin Urol.* 2024;65(3):199-201. doi:[10.4111/icu.20240015](https://doi.org/10.4111/icu.20240015)
 23. Kimmerle J, Timm J, Festl-Wietek T, Cress U, Herrmann-Werner A. Medical students' attitudes toward AI in medicine and their expectations for medical education. *J Med Educ Curric Dev.* 2023;10:1-3. doi:[10.1177/23821205231219346](https://doi.org/10.1177/23821205231219346)
 24. Federspiel F, Mitchell R, Asokan A, Umana C, McCoy D. Threats by artificial intelligence to human health and human existence. *BMJ Glob Health.* 2023;8(5):e010435. doi:[10.1136/bmjgh-2022-010435](https://doi.org/10.1136/bmjgh-2022-010435)
 25. Bostrom N. *Superintelligence: Paths, Dangers, Strategies.* Oxford University Press; 2014.
 26. Cussins J. How to Prepare for the Malicious Use of AI. Future of Life Institute. February 28, 2018. Accessed June 15, 2024. <https://futureoflife.org/recent-news/how-to-prepare-for-malicious-ai/>
 27. Kaku M. *The Future of the Mind: The Scientific Quest to Understand, Enhance, and Empower the Mind.* Allen Lane; 2014.
 28. Tai P, Ding P, Wang F, et al. Brain-computer interface paradigms and neural coding. *Front Neurosci.* 2024;17:1345961. doi:[10.3389/fnins.2023.1345961](https://doi.org/10.3389/fnins.2023.1345961)
 29. Jiao Y, Lei M, Zhu J, Chang R, Qu X. Advances in electrode interface materials and modification technologies for braincomputer interfaces. *Biomater Transl.* 2023;4(4):213-233. doi:[10.12336/biomatertransl.2023.04.003](https://doi.org/10.12336/biomatertransl.2023.04.003)
 30. Wang XQ, Sun HQ, Si JY, Lin ZY, Zhai XM, Lu L. Challenges and suggestions of ethical review on clinical research involving brain-computer interfaces. *Chin Med Sci J.* 2024;39(2):131-139. doi:[10.24920/004377](https://doi.org/10.24920/004377)
 31. Savulescu J, Giubilini A, Vandersluis R, Mishra A. Ethics of artificial intelligence in medicine. *Singapore Med J.* 2024;65(3):150-158. doi:[10.4103/singaporemedj.SMJ-2023-279](https://doi.org/10.4103/singaporemedj.SMJ-2023-279)
 32. Basu K, Sinha R, Ong A, Basu T. Artificial Intelligence: how is it changing medical sciences and its future? *Indian J Dermatol.* 2020;65(5):365-370. doi:[10.4103/ijd.IJD_421_20](https://doi.org/10.4103/ijd.IJD_421_20)
 33. After Asilomar. *Nature.* 2015;526(7573):293-294. doi:[10.1038/526293b](https://doi.org/10.1038/526293b)
 34. Brundage M. *Taking superintelligence seriously: Superintelligence: paths, dangers, strategies* by Nick Bostrom (Oxford University Press, 2014). *Futures.* 2015;72:32-35. doi:[10.1016/j.futures.2015.07.009](https://doi.org/10.1016/j.futures.2015.07.009)
 35. Pfeffer MA, Ling SSH, Wong JKW. Exploring the frontier: transformer-based models in EEG signal analysis for braincomputer interfaces. *Comput Biol Med.* 2024;178:108705. doi:[10.1016/j.compbiomed.2024.108705](https://doi.org/10.1016/j.compbiomed.2024.108705)
 36. Vilhekar RS, Rawekar A. Artificial Intelligence in genetics. *Cureus.* 2024;16(1):e52035. doi:[10.7759/cureus.52035](https://doi.org/10.7759/cureus.52035)
 37. Zhang J, Zhang ZM. Ethics and governance of trustworthy medical artificial intelligence. *BMC Med Inform Decis Mak.* 2023;23(1):7. doi:[10.1186/s12911-023-02103-9](https://doi.org/10.1186/s12911-023-02103-9)
 38. Siafakas NM. Do we need a Hippocratic Oath for artificial intelligence scientists? *AI Magazine.* 2021;42(4):57-61. doi:[10.1609/aaai.12022](https://doi.org/10.1609/aaai.12022)
 39. Subasi ID, Özçelik ŞB. Artificial Intelligence in Breast Imaging: Opportunities, Challenges, and Legal-Ethical Considerations. *Eurasian J Med.* 2023;55(1):114-119. doi:[10.5152/eurasianjmed.2023.23360](https://doi.org/10.5152/eurasianjmed.2023.23360)
 40. Ethics and governance of artificial intelligence for health: WHO guidance. WHO. June 28, 2021. Accessed June 15, 2024. <https://www.who.int/publications/i/item/9789240029200>
 41. The ACT Texts. EU Artificial Intelligence Act. 2024. Accessed June 15, 2024. <https://artificialintelligenceact.eu/the-act/>