Editorial

THE RISKS OF RELYING ON ARTIFICIAL INTELLIGENCE (AI) AS A PRIMARY SOURCE OF HEALTH INFORMATION

Ahmad M [1], Rashid MH[2]

Abstract
As artificial intelligence becomes increasingly integrated into healthcare systems, its role as a primary source of medical information raises ethical and clinical concerns. From symptom checkers to mental health bots and many people even turn to AI as a source of first-line information. The shift, although is a forward step towards healthcare, it also carries serious risks when AI substitutes professional care rather than supplementing it. Accumulating recent empirical and review literature, this editorial examines the interconnected concerns associated with such as bias and overgeneralization, depersonalization of care, misinformation, particularly for vulnerable populations, over dependence on low-cost AI alternatives, and the potential for harm via misdiagnosis, delay of diagnosis/treatment and avoidance of professional treatment. The evidence underscores that AI must remain a tool under human oversight, not an alternative for clinicians.

Introduction

AI's advance into healthcare has created a promise and peril, considering it can democratize access to medical information, especially where clinical access is limited but it can also displace nuanced judgement. The allure for AI lies in the convenience and affordability even though the ethical and social costs are undermined. As patients growingly turn to AI platforms for their health concerns, the risks increase when these sources become a trusted primary source of information. To safeguard patient care and welfare, it is important that we explore the risks as well.

Discussion

AI systems are trained on historical medical and social datasets, which carry embedded biases, whether that be racial, gender, or socioeconomic. Since AI systems learn and accumulate data gathered from academic medical centers, what the systems usually know less about, they will treat with less effectiveness.[1] This can propagate disparities in the care it provides, as algorithms may perpetuate discrimination from pre-existing biases, which can leave affected populations even more marginalized.[2] Parallel to that, AI lacks critical decision-making processes and are often times opaque. Often times called the "black box" effect, which refers to AI's inability to explain its reasoning and logic.[3] Furthermore, AI requires massive amounts of sensitive and private data, which increases the risk of privacy breaches and patient data under private custodianship.[4]

Another downside is AI's capacity to generate misinformation, especially for the vulnerable patient population. There have been documented cases of AI bots delivering content that, intentionally or not, encourages body image issues. One of the studies investigating AI's body image effects in the Punjab region of Pakistan found that that AI provided visuals that were negatively correlated with self esteem and body image satisfaction [5] multiple empirical studies show that patients disclose their histories to physicians who are attuned to perceive them empathetically in real time and able to resonate to with the importance of their story [6], something that AI is unable to provide as they lack emotional intelligence and empathy.

1. Dr. Munaiba Ahmad, Intern Doctor, Chittagong Medical College, Chittagong Bangladesh ; Phone: +8801341802529; Email: munaiba.ahmad@gmail.com.Orcid ID: 0009-0008-6846-6399

2. Dr. Md. Humayun Rashid, Assistant Professor, Department of Neurosurgery, East-West Medical College, Dhaka; Phone: +8801717015031; E-mail: dr.humayunsagor@gmail.com; Orcid ID: 0009-0009-6198-1328

Address of Correspondence:Dr. Md. Humayun Rashid, Assistant Professor, Department of Neurosurgery, East-West Medical College, Dhaka; Phone: +8801717015031; E-mail: dr.humayunsagor@gmail.com,Orcid ID: 0009-0008-6846-6399

Just because AI often is more cost-effective compared to professional healthcare, a dangerous over-dependency can emerge. Most people may begin to see AI as a substitute for clinicians not only due to its lack of critical thinking and judgment but also due to its wide accessibility. However, over-reliance on AI can also in turn lead patients to also defer in critical thinking, and when missteps tend to occur, their accountability is blurred.[7]

This entire cascade of bias, opacity, misinformation, over-reliance creates a fertile ground for maltreatment. Patients are more prone to misdiagnosing themselves based on AI advice, delay seeking professional care, and avoiding physicians entirely, believing that AI is a more reliable guide. Some receive not only incorrect, but also incomplete recommendations, which may worsen their health or delay a preventable harm. Ultimately, when AI is used as a form of primary information source, rather than a supplementary tool, the potential for systemic maltreatment doubles.

Artificial intelligence (AI) can synthesize health information rapidly, but overreliance poses safety, ethical, and accuracy concerns. The following summary highlights key risks, evidence, and mitigation strategies supported by recent WHO, FDA, and peer-reviewed literature.

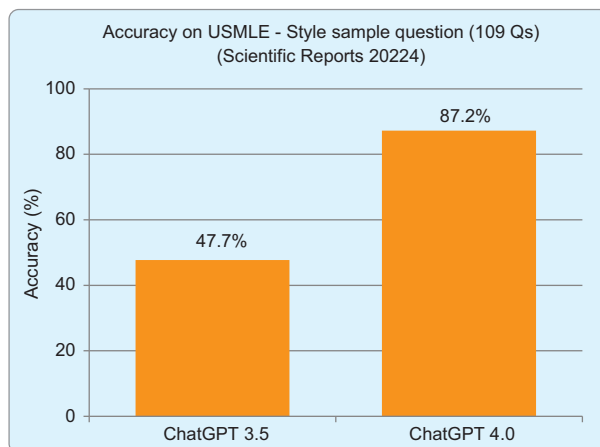| Risk	Evidence / Example	Key Sources	Suggested Mitigation
Hallucinations (false info)	'Basilar ganglia' term in Med-Gemini; literature on LLM hallucinations	The Verge (Med-Gemini); PMC article on AI hallucinations	Human oversight, reference validation
Variable accuracy	ChatGPT 3.5 (47.7%) vs 4.0 (87.2%) USMLE- style Qs; 74.6% correct in top-3 diagnoses	Scientific Reports 2024 [8]	Use validated, task-specific AI; human review
Bias across populations	WHO warns training data bias may cause inequitable outputs	WHO guidance 2024 [9]	Equity audits, diverse datasets
Privacy & regulation gaps	FDA cautions AI/ML SaMD needs transparency	FDA AI/ML guidance 2024–25[10,11,12]	Encryption, PHI restriction, compliance
Automation bias	Clinicians/public may overtrust AI outputs	WHO guidance 2024 [9]	Education, 'AI as decision-support only' policy

Figure 1. *Accuracy on USMLE-style sample questions: ChatGPT 3.5 (47.7%) vs ChatGPT 4.0 (87.2%). Source: Scientific Reports, 2024.*
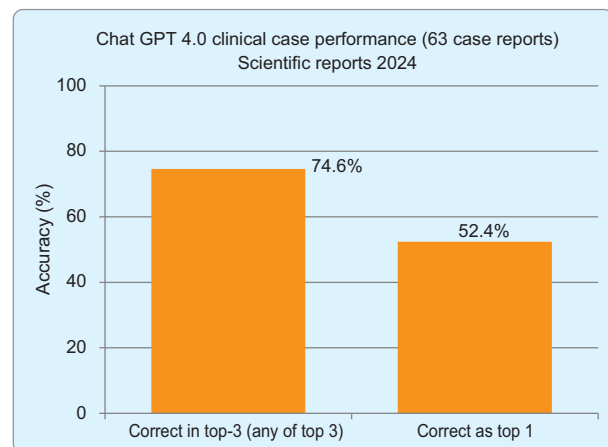

Figure 2. *ChatGPT 4.0 clinical-case performance (63 case reports): correct in top-3 = 74.6%; correct as top-1 = 52.4%. Source: Scientific Reports, 2024.*

Conclusion

AI holds genuine promise to augment healthcare, but its use as a primary source of medical guidance is what carries profound risks if not surveilled. Empirical and review literature relays persistent bias, misinformation, opacity, over-dependence, and harm of AI when it comes to medical advice. These are documented realities that advise us to integrate AI as a complement but not a replacement for professional care of human health. Informed consent, robust regulation, continued monitoring, and accountability is essential. The benefits of AI are possible to harness only while safeguarding the safety, empathetic approach and care that only human clinicians are able to deliver.

References

1. Price WN II. Risks and remedies for artificial intelligence in health care [Internet]. Washington (DC): Brookings Institution; 2019 Nov 14 [cited 2025 Oct 7]. Available from: https://www.brookings.edu/articles/risks-and-remedies-for-artificial-intelligence-in-health-care/

2. Botha NN, Segbedzi CE, Dumahasi VK, Maneen S, Kodom RV, Tsedze IS, Akoto LA, Atsu FS, Lasim OU, Ansah EW. Artificial intelligence in healthcare: a scoping review of perceived threats to patient rights and safety. Arch Public Health. 2024 Oct 23;82(1):188. doi: 10.1186/s13690-024-01414-1. PMID: 39444019; PMCID: PMC11515716.

Peer-Review History:
The peer review history for this paper can be accessed here: https://ewmch.com/review/