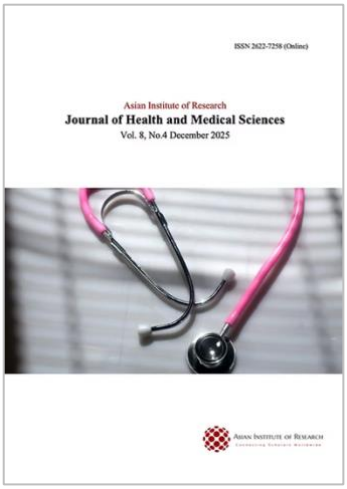

Journal of Health and Medical Sciences

Deonandan, R. (2025), How AI Can Reshape the Fight Against Medical Misinformation. *Journal of Health and Medical Sciences*, 8(4), 9-12.

ISSN 2622-7258

DOI: 10.31014/aior.1994.08.04.244

The online version of this article can be found at:
https://www.asianinstituteofresearch.org/

Published by:
The Asian Institute of Research

The *Journal of Health and Medical Sciences* is an Open Access publication. It may be read, copied, and distributed free of charge according to the conditions of the Creative Commons Attribution 4.0 International license.

The Asian Institute of Research *Journal of Health and Medical Sciences* is a peer-reviewed International Journal. The journal covers scholarly articles in the fields of Medicine and Public Health, including medicine, surgery, ophthalmology, gynecology and obstetrics, psychiatry, anesthesia, pediatrics, orthopedics, microbiology, pathology and laboratory medicine, medical education, research methodology, forensic medicine, medical ethics, community medicine, public health, community health, behavioral health, health policy, health service, health education, health economics, medical ethics, health protection, environmental health, and equity in health. As the journal is Open Access, it ensures high visibility and the increase of citations for all research articles published. The *Journal of Health and Medical Sciences* aims to facilitate scholarly work on recent theoretical and practical aspects of Health and Medical Sciences.

ASIAN INSTITUTE OF RESEARCH
Connecting Scholars Worldwide

The Asian Institute of Research
Journal of Health and Medical Sciences
Vol.8, No.4, 2025: 9-12
ISSN 2622-7258
Copyright © The Author(s). All Rights Reserved
DOI: 10.31014/aior.1994.08.04.244

How AI Can Reshape the Fight Against Medical Misinformation

Raywat Deonandan[1]

[1] Faculty of Health Sciences, University of Ottawa

Correspondence: Raywat Deonandan, Interdisciplinary School of Health Sciences, University of Ottawa, Ottawa, Canada K1H 6N5. E-mail: rdeonand@uottawa.ca

Abstract
Artificial Intelligence (AI) has a significant role in propagating and medical misinformation and disinformation. But it can also be used to mitigate this phenomenon. Natural Language Processing and Sentiment Analysis can be used to analyze vast amounts of data to identify and combat misinformation trends in health-related discussions. AI can be used to categorize the severity of misleading claims, while the analysis and identification of framing strategies can flag misleading content through indirect means. Additionally, AI customizes health messaging to specific audiences to improve engagement and effectiveness. There is also need for enhanced information literacy and regulatory measures to prevent AI misuse, highlighting the dual-edge of AI in modern health communication.

Keywords: Generative AI, Misinformation, ChatGPT

1. Introduction

In the current era, so dependent on digital communication, the rapid dissemination of health-related information frequently includes the spread of medical disinformation. This growing phenomenon poses significant challenges to public health outcomes and represents a pressing crisis. Social media is especially susceptible, as the COVID-19 pandemic revealed. (Wakene et al., 2024) It's not limited to COVID, of course. All health topics are subject to both misinformation and disinformation, from vaccine science to the benefits of pasteurization and the causes of diabetes and heart disease. (Suarez-Lledo & Alvarez-Galvez, 2021) Misinformation is damaging enough. But often, disinformation campaigns are weaponized for political ends, such as for fomenting distrust in the lead-up to an election.

Artificial intelligence (AI) has significantly influenced the spread of medical misinformation, especially through its ability to generate convincing but erroneous content rapidly and at scale. Generative AI models are capable of creating highly realistic text, images, audio, and video that can be indistinguishable from content created by humans, leading to an increase in the volume and sophistication of medical falsehoods online. Intentional mimicry of credible sources is one path by which AI deepens the information crisis. (Monteith et al., 2024) The unintentional offering of erroneous information via AI "hallucinations" is another. (Hatem et al., 2023)

It is a certainty that AI will become increasingly embedded into our lives and economies, and will dictate the avenues and qualities by which we engage in information seeking. Its threats in this regard are well documented. But what of its opportunities? This article explores some of the avenues by which AI can contribute to the improvement of health information sharing, and in combatting the spread of both health misinformation and disinformation.

2. Proactively Identifying Trends in Medical Misinformation

Natural Language Processing (NLP), that aspect of AI that allows machines to interpret, manipulate, and comprehend human language, can be used to analyze vast amounts of textual data from social media, news outlets, and forums where health-related discussions occur. By processing the text, AI can identify recurring themes, new emerging terms, and misinformation patterns. For example, AI can detect a sudden increase in discussions around a specific but unproven treatment method during an infectious disease outbreak.

Sentiment Analysis (SA) evaluates the emotional tone expressed in health communications to gauge public perception around certain medical topics. This can indicate potential areas where misinformation may be taking root if there is a high level of fear or skepticism expressed. Pre-trained models like BERT have been modified to assess emotional states from text. (Hossain et al., 2025) Still in its infancy, this approach can be accelerated through the development of more wide-coverage misinformation datasets, "whose data is multilingual and extracted from a variety of different platforms with varying data formats." (Liu et al., 2024)

Initiatives like "Project Heal" seek to address three types of erroneous online claims: misinformation (things that are factually incorrect), disinformation (intentional errors meant to cause harm) and malinformation (things that are correct but expressed out of context in a misleading fashion). Project Heal trains a large language model to categorize likely sources of such incorrect claims, and attempts to rate them by severity and potential impact. (Rama, 2024) A flotilla of tools incorporating Project Heal's ranking approach, NLP, and SA will doubtless prove invaluable in the deepening real-time battle against misinformation.

3. Identifying Framing Strategies

Rather than arbitrating whether a statement is true or false, AI can assess the rhetorical strategies used in a given claim. In other words, AI can identify a statement's "frame." The four identifiable elements of a communication frame in a piece of health text or a snippet of video or audio are its problem definition, causal interpretation, moral evaluation, and treatment recommendation. (Entman, 1993) Savvy communicators know that the choice of frame makes information more noticeable, meaningful, and memorable. (Entman, 1993)

Consider a selection of news articles about the same topic, such as gun violence. One might emphasize gun control, while another might promote gun ownership rights, and a third might emphasize mental health issues. (Liu et al., 2019) In the realm of public health, a given article might emphasize uncertainty in safety around COVID-19 vaccines, or it might try to establish legitimacy by referencing an authority. (Sepulvado & Burke-Garcia, 2024) These are all framing strategies.

The Framed Element-based Model (FEM) proposed by Wang et al (Wang et al., 2024) seeks to use AI to detect likely misinformation in news articles by assessing problem definition, claims of causality, moral positioning, and whether a preferred treatment is recommended. In this way, the AI ascertains falseness by establishing framing patterns that have been observed to be preferred by disinformation and malinformation merchants. By identifying the framing strategy, one does not have to determine the correctness of a claim, but rather only the way that the claim was positioned and communicated.

4. Crafting Health Messaging for Specific Audiences

By personalizing the delivery of health information, AI can help ensure that accurate messages are more engaging and reach a wider audience. In the words of Andrew Beam, "One benefit of [AI models] is they are very good at

modulating their voice. They can meet people where they are, delivering high-quality information that's both easy to understand and accessible to folks from a wide variety of demographic and cultural backgrounds." (Beam, 2024) AI can segment audiences based on demographics, health conditions, and online behaviour, among other factors. By identifying subgroups within a larger population, AI can help create messages that are specifically tailored to the characteristics and needs of each group. AI algorithms can analyze past interactions to determine which types of messages and delivery methods have been most effective for different population segments in the past. This information can be used to customize messages by tone, complexity, and format.

Additionally, AI can predict the health information needs of different individuals or groups based on their health trajectories or risk profiles. One can envision an AI model that can continuously learn from how different audiences respond to various health messages. Such a feedback loop would allow for the refinement and optimization of communication, raising the probability that messaging remains effective over time.

5. Final Thoughts

In the digital age, the challenge of medical misinformation has become a significant public health concern, intensified by the capabilities of artificial intelligence to both propagate and combat such misinformation. It is clear that the proactive use of AI will be indispensable and probably unavoidable. But the technology cannot be relied upon to manage this task in isolation. The promotion of both information literacy and transparency in AI training datasets is vital. (Germani et al., 2024) Regrettably, there may eventually be a need to enact laws to prevent the weaponization of AI in the production of medical falsehoods. (Haupt & Marks, 2024).

Funding: This project is part of the funded Chair in University Teaching at the University of Ottawa.

Conflict of Interest: The authors declare no conflict of interest.

Informed Consent Statement/Ethics Approval: Not applicable.

Declaration of Generative AI and AI-assisted Technologies: This study has not used any generative AI tools or technologies in the preparation of this manuscript.

References

Beam, A. (2024). Misinformation doesn't have to get the last word. Harvard Public Health. https://harvardpublichealth.org/tech-innovation/how-ai-misinformation-could-impact-the-future-of-public-health/

Entman, R. (1993). Framing: Toward Clarification of A Fractured Paradigm. The Journal of Communication, 43, 51-58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Germani, F., Spitale, G., & Biller-Andorno, N. (2024). The Dual Nature of AI in Information Dissemination: Ethical Considerations. Jmir ai, 3, e53505. https://doi.org/10.2196/53505

Hatem, R., Simmons, B., & Thornton, J. E. (2023). A Call to Address AI "Hallucinations" and How Healthcare Professionals Can Mitigate Their Risks. Cureus, 15(9), e44720. https://doi.org/10.7759/cureus.44720

Haupt, C. E., & Marks, M. (2024). FTC Regulation of AI-Generated Medical Disinformation. Jama, 332(23), 1975-1976. https://doi.org/10.1001/jama.2024.19971

Hossain, M. M., Hossain, M. S., Mridha, M. F., Safran, M., & Alfarhood, S. (2025). Multi task opinion enhanced hybrid BERT model for mental health analysis. Sci Rep, 15(1), 3332. https://doi.org/10.1038/s41598-025-86124-6

Liu, S., Guo, L., Mays, K., Betke, M., & Wijaya, D. (2019). Detecting Frames in News Headlines and Its Application to Analyzing News Framing Trends Surrounding U.S. Gun Violence. https://doi.org/10.18653/v1/K19-1047

Liu, Z., Zhang, T., Yang, K., Thompson, P., Yu, Z., & Ananiadou, S. (2024). Emotion detection for misinformation: A review. Information Fusion, 107, 102300. https://doi.org/https://doi.org/10.1016/j.inffus.2024.102300

Monteith, S., Glenn, T., Geddes, J. R., Whybrow, P. C., Achtyes, E., & Bauer, M. (2024). Artificial intelligence and increasing misinformation. The British Journal of Psychiatry, 224(2), 33-35. https://doi.org/10.1192/bjp.2023.136

Rama, G. (2024). AWS, Researchers Developing AI To Fight Medical Misinformation. AWS Insider. https://awsinsider.net/Articles/2024/06/03/AWS-Researchers-AI-Medical-Misinfo.aspx

Sepulvado, B., & Burke-Garcia, A. (2024). AI and Misinformation on Social Media: Addressing Issues of Bias and Equity across the Research-to-Deployment Process. American Association for Public Opinion Research. https://aapor.org/newsletters/ai-and-misinformation-on-social-media-addressing-issues-of-bias-and-equity-across-the-research-to-deployment-process/

Suarez-Lledo, V., & Alvarez-Galvez, J. (2021). Prevalence of Health Misinformation on Social Media: Systematic Review. J Med Internet Res, 23(1), e17187. https://doi.org/10.2196/17187

Wakene, A. D., Cooper, L. N., Hanna, J. J., Perl, T. M., Lehmann, C. U., & Medford, R. J. (2024). A pandemic of COVID-19 mis- and disinformation: manual and automatic topic analysis of the literature. Antimicrob Steward Healthc Epidemiol, 4(1), e141. https://doi.org/10.1017/ash.2024.379

Wang, G.-H., Frederick, R., Duan, J., Wong, W., Rupar, V., Li, W., & Bai, Q.-w. (2024). Detecting misinformation through Framing Theory: the Frame Element-based Model. ArXiv, abs/2402.15525.