Journal of Academic Inquiry and Knowledge Volume 2023 (1) 43-50

THE ETHICAL CONCERNS OF ARTIFICIAL INTELLIGENCE MODELS IN HEALTH AND MEDICINE

Jayden Choi

Los Angeles, CA, USA

The Ethical Concerns of Artificial Intelligence Models in Health and Medicine

The release of ChatGPT in late November 2022 brought renewed attention to artificial intelligence's potential for application in many fields. One of the fields that artificial intelligence (AI) can be of great use is healthcare. Medicine has seen significant integration of AI. Deep-learning AI models diagnose illnesses, track health patterns, and advance treatment (Mozafaripour, 2020). For example, the National Cancer Institute designed an AI algorithm that consistently outperformed human experts in recognizing cervical cancer (*Researchers Create AI Approach for Cervical Cancer Screening - NCI*, 2019).

AI has made its mark on the healthcare industry and promises to grow. In 2019, AI spending in the healthcare and pharmaceutical industries was estimated to be around $463 million and is predicted to grow to $2 billion by 2025 (*COVID-19 Pandemic Impact*, n.d.). However, as with any technology, potential issues must be considered. First, as AI uses comprehensive datasets to determine diagnoses (*Be Part of Research*, n.d.), these models must be trusted with the sensitive information of millions of patients, many of whom may not have provided consent. Furthermore, AI models are capable of expressing bias in their conclusions.

These ethical issues can lead to troubling situations. Therefore, while the applications of AI in healthcare could lead to considerable advancements in medicine and wellness, ethical concerns of AI application in healthcare must be considered, primarily informed consent, data privacy, and bias.

Concerns of Informed Consent

The implications of informed consent in using AI in healthcare are paramount. Informed consent is a necessary aspect of healthcare and requires an "assessment of the patient's understanding" of the entire medical procedure (Shah et al., 2023, p. 1). This includes all information regarding the usage and collection of the patient's medical data. Due to the nature of AI and its machine learning model, data collection is a vital aspect of training AI models (Becanovic, 2021). Although obtaining patients' consent for data collection is not a significant concern, the issue lies in the patient's understanding of their data usage.

As AI models use complex algorithms, it may be difficult for human healthcare providers to sufficiently understand AI procedures (Astromskė et al., 2021). This issue is widely discussed within AI research communities and has been dubbed the "black box effect" by researchers. The *black box effect* can be defined as "an artificial intelligence (AI) system, device, or program" which can provide "useful information without revealing any information about [the system's] internal workings," thus leaving the understanding of its internal mechanisms. ("The Black Box Effect - How Can AI and ML Provide Transparent Insights for Drug Discovery?" n.d.).

The *black box effect* ties into the concerns of informed consent, as this situation would not allow patients a comprehensive understanding of the exact impacts of AI on their healthcare procedures. For example, while patients may be able to understand that their data may (with their consent) be used to train AI models, they may not know how it may be used, as the AI

algorithms are too complex to be explained. Likewise, a healthcare worker may be able to convey to a patient that their medical procedure may be assisted by AI, but may not be able to convey the workings of AI in the procedure.

While the issue of informed consent is a major limitation of integrating AI in healthcare, potential solutions exist. One solution proposes that the burden of explanation could be shifted from the healthcare workers to the developers of AI systems (Astromskė et al., 2021). This solution would maintain the legal integrity of informed consent, as the AI developers could explain AI's methods to patients. This solution would still require healthcare workers to understand the basic concepts of the integration of AI in medical procedures but allows AI experts to explain the more complex aspects of AI involvement.

Concerns of Patient Data Privacy and Usage

Another major issue of the integration of AI in healthcare lies in the privacy of patients' data. As mentioned, patients cannot be guaranteed that their data will be kept private, and many will not know how it is used (Li et al., 2023). Again, AI models are trained using vast amounts of recorded data (Glaser & Niebla, 2023) which could include data of private patients who may not want their data used in AI training. This ties back into informed consent, as patients may not have been able to consent to the contribution of their data to AI training. Furthermore, this data is also at risk of data breaches. While data breaches remain a problem in almost all technologies, they are especially problematic in AI due to the large amounts of data that could be compromised (Glaser & Niebla, 2023). AI data can potentially compromise far more information than other data breaches.

Likewise, unethical uses of data usage involving AI are also of concern, such as the prospect of financial benefit. For example, a 2023 study showed how insurance companies might

be able to use AI algorithms to determine whether or not to approve claims (Li et al., 2023). While this may lead to financial benefits for the insurance companies, it is also clearly unethical as it limits the benefits of insurance healthcare based on monetary concerns. As the issues of data privacy and usage issues prove to be a major concern in integrating AI within healthcare, it is important that solutions be considered. For one, it must be ensured that regulations regarding data security and privacy involving AI are specific and heavily regulated. Furthermore, another solution is the option for patients to opt out of certain types of data collection (Bartneck et al., 2021).

Concerns of Bias in Healthcare AI Models

Finally, the possibility of bias in AI systems has also been demonstrated as a significant concern in incorporating AI in healthcare. For example Morley et al. discuss the possibility of demographic discrimination in a 2020 study. They posit that if an AI model is trained using data primarily representing one demographic group, it may create patterns that are only relevant to that specific demographic group (2020). This bias can lead to discrimination when the AI model treats other demographic groups (Obermeyer et al., 2019). For example, suppose an AI model is primarily trained using medical data of Caucasian patients, who may present specific patterns based on a specific characteristic. In that case, it may incorrectly assume that these patterns exist in the treatment of an Asian patient, leading to issues such as misdiagnosis.

Furthermore, the possibility of economic discrimination also exists in AI systems. Similar to demographic discrimination in AI, economic discrimination can occur if AI models are not trained sufficiently to recognize this. As populations from different economic backgrounds present differing lifestyle conditions, AI models trained primarily using data from patients of a specific economic background may provide inaccurate information for patients that do not match

this economic background. Economic discrimination may also occur in treatments and prescriptions involving AI. If, again, an AI model is based on data primarily from patients sharing one economic background, it may reflect discriminatory behavior by recommending treatment solutions and medications that simply may not be viable for patients of another economic background.

In order to combat this issue, AI models must specifically be trained to recognize possible bias and discrimination in certain groups. Leslie et al. recommend that "clinical expertise, inclusive community involvement, interdisciplinary knowledge, and ethical reflexivity" be incorporated into AI training to ensure that demographic and economic discrimination is limited (2021).

Conclusion

There are serious concerns regarding the integration of AI in healthcare. Concerns of informed consent, data privacy, and bias have all been proven to be troubling matters in healthcare and AI, especially as AI is increasingly integrated into health and medicine. As AI is such an advanced and unfamiliar technology, these issues seem especially daunting for many. However, with appropriate vigilance and precaution, such as the solutions described above, progress can be made to harness the potential of AI to benefit the world of medicine.

References

Astromskė, K., Peičius, E., & Astromskis, P. (2021). Ethical and legal challenges of informed consent applying artificial intelligence in medical diagnostic consultations. *AI & SOCIETY*, *36*(2), 509–520. https://doi.org/10.1007/s00146-020-01008-9

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). Privacy Issues of AI. In C. Bartneck, C.Lütge, A. Wagner, & S. Welsh, *An Introduction to Ethics in Robotics and AI* (pp. 61–70).Springer International Publishing. https://doi.org/10.1007/978-3-030-51110-4_8

Be Part of Research. (n.d.). Retrieved April 25, 2023, from https://bepartofresearch.nihr.ac.uk/articles/artificial-intelligence/

Becanovic, S. (2021, December 13). What is training data? *StageZero Technologies*. https://stagezero.ai/blog/what-is-training-data/

COVID-19 Pandemic Impact: Global R&D Spend For AI in Healthcare and Pharmaceuticals Will Increase US$1.5 Billion By 2025. (n.d.). Retrieved April 25, 2023, from https://www.abiresearch.com/press/covid-19-pandemic-impact-global-rd-spend-ai-healthcare-and-pharmaceuticals-will-increase-us15-billion-2025/

Glaser, J. D., & Niebla, L. (2023, April 5). *Benefits and Legal Risks of Embracing Generative AI Applications | Mintz*. https://www.mintz.com/insights-center/viewpoints/2911/2023-04-05-benefits-and-legal-risks-embracing-generative-ai

Leslie, D., Mazumder, A., Peppin, A., Wolters, M. K., & Hagerty, A. (2021). Does "AI" stand for augmenting inequality in the era of covid-19 healthcare? *BMJ*, *372*, n304. https://doi.org/10.1136/bmj.n304

Li, F., Ruijs, N., & Lu, Y. (2023). Ethics & AI: A Systematic Review on Ethical Concerns and Related Strategies for Designing with AI in Healthcare. *AI*, *4*(1), Article 1. https://doi.org/10.3390/ai4010003

Morley, J., Machado, C. C. V., Burr, C., Cowls, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review. *Social Science & Medicine*, *260*,113172. https://doi.org/10.1016/j.socscimed.2020.113172

Mozafaripour, S. (2020, June 10). How AI Is Revolutionizing Healthcare | USAHS. *University of St. Augustine for Health Sciences*. https://www.usa.edu/blog/how-ai-is-revolutionizing-healthcare/

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Researchers create AI approach for cervical cancer screening—NCI (nciglobal,ncienterprise). (2019, January 10). [CgvPressRelease]. https://www.cancer.gov/news-events/press-releases/2019/deep-learning-cervical-cancer-screening

Shah, P., Thornton, I., Turrin, D., & Hipskind, J. E. (2023). Informed Consent. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK430827/

The Black Box Effect—How Can AI and ML Provide Transparent Insights for Drug Discovery? (n.d.). *Oxford Global*. Retrieved April 16, 2023, from https://www.oxfordglobal.co.uk/resources/the-black-box-effect/

Wang, C., Zhang, J., Lassi, N., & Zhang, X. (2022). Privacy Protection in Using Artificial Intelligence for Healthcare: Chinese Regulation in Comparative Perspective. *Healthcare*, *10*(10), Article 10. https://doi.org/10.3390/healthcare10101878