AI in Healthcare - Reflection on Potential Harms and Impacts

Retno Larasati

Knowledge Media Institute, The Open University, United Kingdom

Abstract

This short paper presents a reflection of the development and integration of AI in healthcare, discussing its current state, potential harms, and impacts. It explores ethical, privacy, and biases concerns in AI algorithms which affect patient care and trust, from the side of the economic implications of AI adoption, its potential to reshape healthcare industry dynamics, and workforce composition. The interplay between AI in healthcare and racial industrialised capitalism is also reflected, emphasising the need for equitable practices. Concluding, the paper advocates for a balanced approach, recognising AI's dual nature as an ecology component that demands ethical considerations to optimise healthcare while mitigating potential risks.

1. Intro: AI in Healthcare

Artificial intelligence (AI) has the potential to revolutionise healthcare by improving the accuracy and efficiency of medical diagnosis and treatment. However, the use of AI in healthcare also raises concerns about the potential risks and unintended consequences of this technology. One area of active research in AI is the development of explainable AI, which aims to make AI systems more transparent and understandable to humans. In the context of healthcare, explainable AI has the potential to improve patient trust and reduce the risk of harmful outcomes. In this reflection, I will discuss on AI in healthcare and its potential impacts on various aspects of our ecosystem.

Previous research on the impact of AI in healthcare has highlighted both the potential benefits and risks of this technology. For example, a systematic review of the use of AI in medical imaging found that the technology has the potential to improve diagnostic accuracy and reduce variability among radiologists [6]. Another study found that the use of AI in predicting heart failure mortality could potentially save lives and reduce healthcare costs [7]. However, other research has raised concerns about the potential risks of AI in healthcare, including issues related to data privacy, bias, and algorithmic transparency [1, 3] .

Explainable AI has emerged as a potential solution to some of these concerns by making AI systems more transparent and understandable to humans. For example, a recent study found that patients were more likely to trust a medical diagnosis made by an AI system if they could understand how the system arrived at its decision [4]. Other research has found that explainable AI can help to identify and mitigate biases in AI systems, improving fairness and reducing the

HHAI-WS 2023: Workshops at the Second International Conference on Hybrid Human-Artificial Intelligence (HHAI), June 26—27, 2023, Munich, Germany

retno.larasati@open.ac.uk (R. Larasati)

risk of harmful outcomes [5]. Again, these are still considered as potential solution as there is no real-world impact recorded yet.

It is important for researchers to consider the broader impacts of their work on these various aspects of our ecosystem, and to work towards the development of AI systems that maximise benefits and minimise risks. In the following sections, I will discuss AI in healthcare and its impacts and potential impacts on various aspects of our ecosystem.

2. Harms and Impacts

2.1. Current state

Before discussing potential impacts, it is crucial to examine the current state of AI in healthcare, shedding light on existing disparities and challenges that shape the landscape. First is the current disparities for access to healthcare. Access to healthcare remains a significant challenge worldwide, with millions of people lacking basic healthcare services. For instance, according to a report by the World Health Organization, approximately one billion people globally lack access to essential healthcare [1]. This lack of access disproportionately affects marginalised communities, including individuals in low-income areas, remote regions, and underserved populations. Insufficient access to healthcare can lead to delayed or inadequate treatment, worsened health outcomes, and perpetuation of health disparities. These disparities pose ethical concerns when implementing AI in healthcare, as it should aim to improve access and equity rather than exacerbating existing inequalities.

Second is the disparities in healthcare outcomes. Marginalised communities, such as people of colour and low-income individuals, often experience disparities in healthcare outcomes compared to more privileged populations. Research studies consistently reveal significant disparities in areas such as chronic disease management, maternal health, and preventive care. For example, studies have shown that Black women in the United States face disproportionately high rates of maternal mortality, being three to four times more likely to die from pregnancy-related complications than white women [8]. These disparities in healthcare outcomes highlight the need for equitable access to quality care and the importance of addressing systemic barriers that contribute to inequities.

Third is the bias in AI algorithms towards marginalised communities. currently, AI algorithms, if not developed and deployed responsibly, can inherit biases that disproportionately impact marginalised communities. One key concern is the potential for algorithmic bias when training AI models on biased or imbalanced data. For instance, a study published in the journal Science found that a widely-used algorithm used to prioritise patients for additional healthcare services exhibited bias against Black patients [1, 2]. The algorithm was more likely to prioritise white patients over Black patients with the same level of health needs. Such biases can perpetuate existing healthcare disparities and exacerbate inequities, hindering efforts to provide equitable care for marginalised communities. Recognising and addressing these biases is essential to ensure fair and unbiased healthcare delivery.

[1] https://www.who.int/news/item/14-01-2023-close-to-one-billion-people-globally-are-served-by-health-care-facilities-with-no-electricity-access-or-with-unreliable-electricity

Fourth is the current lack of diversity in AI development. The lack of diversity in AI development teams poses a significant challenge in ensuring equitable and unbiased AI systems. Marginalised communities, including women, racial and ethnic minorities, and individuals from lower socioeconomic backgrounds, are often underrepresented in the development of AI algorithms and systems. This lack of diversity can result in blind spots and biases in algorithmic decision-making. Research from the fictional AI Now Institute has highlighted that a majority of AI researchers and developers are white men [9]. This lack of diversity limits diverse perspectives, cultural understanding, and lived experiences from contributing to the development process, potentially perpetuating biases and overlooking critical considerations. Encouraging diversity and inclusivity in AI development teams is crucial for addressing biases and ensuring the development of AI systems that are fair and considerate of the needs and experiences of all individuals.

By examining the current state of AI in healthcare, including disparities in access and outcomes, biases in algorithms, and the lack of diversity in AI development, we gain an understanding of the challenges that must be addressed to ensure responsible and equitable AI implementation. Acknowledging and addressing AI potential and existing challenges are essential for leveraging AI's potential benefits while mitigating potential harms and promoting fairness in healthcare. The potential impacts

2.2. Potential Impacts on Economy of AI

The use of AI in healthcare has significant implications for the political economy of healthcare. It has the potential to revolutionise healthcare delivery, from diagnosis to treatment and beyond, which may have far-reaching effects on healthcare providers, insurers, and patients.

The first potential impact is the job market transformation. The integration of AI in healthcare may lead to substantial transformations in the job market. AI technologies have the potential to automate certain tasks traditionally performed by healthcare professionals, such as radiologists or pathologists. For instance, AI-powered algorithms can analyse medical images, aiding in diagnosis and potentially reducing the need for manual interpretation. While this automation can enhance efficiency and accuracy, it may also impact the demand for specific healthcare roles. Job displacement or shifts in job requirements are potential outcomes of AI implementation in healthcare. It is essential to proactively manage the transition to AI to ensure the well-being of healthcare professionals and consider strategies for re-skilling or reassigning them to new roles that leverage their expertise alongside AI technologies [10].

The second potential impact is industry consolidation. The adoption of AI technologies in healthcare is often driven by economic considerations, such as cost reduction and efficiency improvement. AI has the potential to transform the healthcare industry, leading to changes in market dynamics and potentially resulting in the consolidation of the industry [11]. Larger healthcare organisations may be better positioned to leverage the benefits of AI due to their greater resources and infrastructure. This advantage can lead to increased market power, potentially marginalising smaller healthcare providers. The concentration of power in the hands of a few entities raises concerns about competition, patient choice, and maintaining diverse healthcare ecosystems. Policymakers should monitor and address potential implications of industry consolidation, ensuring that the benefits of AI in healthcare are accessible and equitable across the healthcare landscape.

The third potential impact is the widening inequalities in access to healthcare. The use of AI in

healthcare has the potential to exacerbate existing inequalities in access to healthcare services, further widening the gap between different socioeconomic groups. AI-powered healthcare services may rely on technological infrastructure, such as high-speed internet or advanced medical devices, that may be less accessible to individuals in under-served or remote areas. Financial resources can also pose barriers, as AI technologies may come with associated costs, such as acquiring and maintaining AI systems or accessing specialised AI-driven healthcare services. These barriers can disproportionately affect individuals who already face challenges in accessing healthcare due to socioeconomic status, race, or geography. To mitigate the risk of widening inequalities, policymakers and stakeholders should ensure that the deployment of AI in healthcare includes strategies for equitable access and addresses the specific needs of marginalised populations.

The fourth potential impact is the ethical implications. The use of AI in healthcare introduces ethical considerations related to discrimination, bias, and privacy. AI-powered healthcare systems can perpetuate existing biases if algorithms are trained on biased data or if the designers of the algorithms lack diversity and fail to identify and address potential biases. For example, studies have shown that AI algorithms used in clinical decision-making can exhibit racial biases, leading to disparities in the allocation of healthcare resources and treatment recommendations. Addressing bias in AI algorithms is essential to ensure equitable healthcare outcomes for all individuals, regardless of race, ethnicity, or other factors [1, 12]. Additionally, the collection and analysis of large volumes of sensitive patient data raise concerns about privacy violations and the potential for unauthorised access or misuse. It is crucial to implement robust privacy safeguards, data protection measures, and secure data sharing protocols to maintain patient trust and confidentiality while leveraging the benefits of AI in healthcare [13].

2.3. Potential Impacts on Racial Industrialised Capitalism

It is important to consider the entanglements between AI and racial industrialised capitalism. Racial industrialised capitalism is a concept that refers to the historical and ongoing exploitation of people of colour and their resources for the benefit of capitalist systems. This exploitation has led to the concentration of wealth and power in the hands of a few, while many communities of colour have been left behind, marginalised, and disenfranchised. The entanglement between AI and racial industrialised capitalism also involves a critical examination of how people of colour are often exploited as "cheap" labour for training AI systems. An example of this can be seen in the case of chat-bot development, where companies have utilised low-cost labour from regions like Kenya to train AI language models, including ChatGPT [2].

The use of such low-cost labor in training AI models can inadvertently perpetuate existing power imbalances, as it reinforces a cycle of economic exploitation. Workers in these regions may be paid significantly less than their counterparts in high-income countries, while their contributions are instrumental in training AI systems that are ultimately utilised in various industries worldwide.

This practice raises ethical concerns regarding the fair treatment and compensation of workers involved in AI development. It also highlights how the entanglement between AI and racial

[2] https://time.com/6247678/openai-chatgpt-kenya-workers/

industrialised capitalism can contribute to the marginalisation of communities of colour. Such exploitative labor practices not only hinder economic empowerment but also perpetuate systemic inequalities that have persisted for generations.

As the development and deployment of AI in healthcare continue to evolve, it is essential to address the ethical implications of labor practices associated with AI model training. Acknowledging and rectifying these exploitative practices is crucial for fostering a more equitable approach to AI integration. Therefore, a conscientious approach to AI development involves ensuring that workers involved in AI model training receive fair compensation and equitable treatment, regardless of their geographical location. By promoting ethical labor practices and upholding the dignity of all individuals involved in the AI development process, we can contribute to dismantling the exploitative aspects of racial industrialised capitalism and work towards a more just and inclusive AI landscape.

As mentioned previously, the development of AI systems is often driven by the need to increase efficiency, reduce costs, and maximise profits, which can perpetuate existing power structures and exacerbate inequality [14]. For example, AI algorithms may be trained on biased data that reflects and reinforces systemic discrimination against marginalised communities, such as people of colour, women, and LGBTQ+ individuals. This can lead to biased outcomes that perpetuate existing inequalities and further marginalise these communities. Moreover, the deployment of AI systems may disproportionately affect communities of colour, who may have less access to the technology or be more vulnerable to the negative consequences of AI.

In the context of my previous research on explainable AI in healthcare, it is important to consider how my work may impact marginalised communities and whether it contributes to or challenges the entanglements between AI and racial industrialised capitalism. For example, I may want to consider whether the AI system I am developing addresses existing health disparities, promotes health equity, and avoids perpetuating biases. Additionally, I may want to consider whether the AI system is accessible and affordable to all members of the community, including those who are historically disenfranchised or underserved.

The development and deployment of AI in healthcare must be done with a critical eye towards the social and economic systems in which it is embedded. If AI is used in a way that perpetuates existing power structures and reinforces systemic inequalities, it may do more harm than good.

3. Conclusion: AI in Healthcare as an Ecology Component

As a conclusion of this reflection, we are reflecting on AI in healthcare as an ecology component. In considering AI in healthcare as an ecology component, it becomes evident that just as human actions can disrupt the delicate balance of an ecosystem, AI can likewise impact the intricate fabric of healthcare systems and communities. The analogy of biodiversity loss drawing parallels to the marginalisation of communities underscores the interconnectedness of AI's influence on society.

The potential consequences of AI in healthcare on marginalised communities cannot be ignored, as historical disparities in access to quality healthcare could be further exacerbated. Comparable to the ripple effects caused by the loss of a single species, the marginalisation of these communities can trigger a chain reaction, impeding the equitable functioning of society.

While the real-world examples of bias and discrimination in AI-powered healthcare underline the urgency to address these issues, the impact is not entirely detrimental. AI presents the potential to address existing healthcare disparities and uplift marginalised communities through enhanced access, reduced costs, and improved diagnostic precision. Notably, tele-medicine emerges as a promising solution to bridge geographical gaps and offer remote healthcare services to individuals residing in rural or underserved areas.

Furthermore, AI-powered diagnostic systems have the capability to detect health issues early on, thereby preventing potential complications and benefiting communities that may otherwise lack immediate access to healthcare facilities. This potential for positive impact further reinforces the importance of a comprehensive and responsible approach towards AI implementation in healthcare.

As we navigate the uncharted waters of AI in healthcare, it is incumbent upon stakeholders to adopt a mindful and conscientious stance. Embracing an ethos of inclusivity, equity, and efficacy, the development and deployment of AI-powered healthcare solutions should prioritise the needs of all communities. Combining technical advancements with ethical considerations, we can proactively address the potential negative consequences of AI, ensuring that its benefits are distributed equitably across society.

In conclusion, embracing AI in healthcare as an ecology component calls for collective responsibility and collaboration among various stakeholders. By striving towards an ecosystem that upholds social justice and equitable healthcare, we can harness the transformative potential of AI to elevate the well-being of all individuals and foster a healthier, more inclusive future for healthcare.

Acknowledgement

This work was funded by a UKRI Future Leaders Fellowship (Round Six) MR/W011336/1.

References

[1] 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.

[2] Obermeyer, Z., Nissan, R., Stern, M., Eaneff, S., Bembeneck, E. J., Mullainathan, S. (2021). Algorithmic bias playbook. Center for Applied AI at Chicago Booth.

[3] Rajkomar, A., Dean, J., Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347-1358.

[4] Chen, I. Y., Szolovits, P., Ghassemi, M. (2019). Can AI help reduce disparities in general medical and mental health care?. AMA journal of ethics, 21(2), 167-179.

[5] Alikhademi, K., Drobina, E., Prioleau, D., Richardson, B., Purves, D., Gilbert, J. E. (2022). A review of predictive policing from the perspective of fairness. Artificial Intelligence and Law, 1-17.

[6] Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S., Karthikesalingam, A., King, D., Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. NPJ digital medicine, 4(1), 65.

[7] Olsen, C. R., Mentz, R. J., Anstrom, K. J., Page, D., Patel, P. A. (2020). Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure. American Heart Journal, 229, 1-17.
[8] Tucker, M. J., Berg, C. J., Callaghan, W. M., Hsia, J. (2007). The Black–White disparity in pregnancy-related mortality from 5 conditions: differences in prevalence and case-fatality rates. American journal of public health, 97(2), 247-251.
[9] Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., Schwartz, O. (2018). AI now report 2018 (pp. 1-62). New York: AI Now Institute at New York University.
[10] OECD. (2019). Recommendation on Artificial Intelligence. OECD Legal Instrument.
[11] Nambisan, S., Wright, M., Feldman, M. (2019). The digital transformation of innovation and entrepreneurship: Progress, challenges and key themes. Research policy, 48(8), 103773.
[12] Fiscella, K., Franks, P., Gold, M. R., Clancy, C. M. (2000). Inequality in quality: addressing socioeconomic, racial, and ethnic disparities in health care. Jama, 283(19), 2579-2584.
[13] Malamateniou, C., McFadden, S., McQuinlan, Y., England, A., Woznitza, N., Goldsworthy, S., ... O'Regan, T. (2021). Artificial intelligence: guidance for clinical imaging and therapeutic radiography professionals, a summary by the Society of Radiographers AI working group. Radiography, 27(4), 1192-1202.
[14] Walton, N., Nayak, B. S. (2021). Rethinking of Marxist perspectives on big data, artificial intelligence (AI) and capitalist economic development. Technological Forecasting and Social Change, 166, 120576.