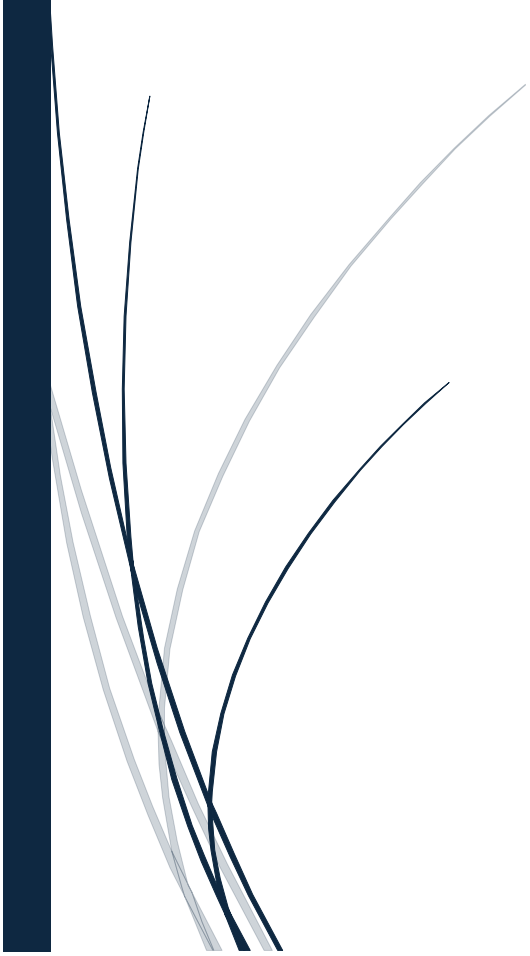
6/17/2025

Initials and Surname: EC Terblanche

Student Number: 3020408

FAMH4007A- Innovative health

Podcast and Report



Ethan Terblanche
3020408 Podcast and Report

Podcast: Please note that the podcast was submitted with the report as a Mp4

Introduction

This report accompanies a podcast episode that critically explores the ethical issue of algorithmic bias in artificial intelligence (AI) used in healthcare. The discussion focuses on how AI-driven decision-making impacts patient autonomy and justice, examining the implications for diverse stakeholders and proposing practical safeguards to support ethical innovation in healthcare systems. The podcast and report aim to engage a broad audience, including healthcare professionals, policymakers, and the public. By focusing on a current and impactful ethical issue, the project encourages reflection on how emerging technologies intersect with fundamental healthcare values. This approach helps bridge the gap between technical innovation and ethical responsibility.

Rationale

Artificial intelligence is transforming healthcare, from diagnostics and treatment recommendations to resource allocation and patient management. The rapid integration of AI systems has brought significant benefits, such as increased efficiency and improved accuracy in clinical settings. However, these advances have also raised important ethical concerns, particularly regarding fairness, transparency, and accountability (Obermeyer et al., 2019). Algorithmic bias is a pressing issue because it has the potential to reinforce or amplify existing health disparities, especially among underrepresented or marginalised groups (Obermeyer et al., 2019).

This topic was chosen due to its relevance in both local and global contexts, where equitable access to quality healthcare remains a critical goal. In South Africa, for example, diverse patient populations and resource constraints make the ethical deployment of AI particularly significant. Addressing these challenges is essential for fostering trust and ensuring that technological advances benefit all patients, not just a privileged few (Rajkomar et al., 2018).

The increasing reliance on AI in healthcare necessitates a proactive ethical examination to prevent unintended consequences. Algorithmic bias was chosen as the focal point because Obermeyer et al. (2019) highlight how such bias can perpetuate systemic inequalities, making it a critical issue in global health ethics. Moreover, Rajkomar et al. (2018) argue that balancing innovation with equity and social justice remains a central challenge in deploying AI technologies responsibly.

Key Findings

Algorithmic Bias

AI systems in healthcare can inadvertently perpetuate or exacerbate health disparities if trained on non-representative data. For example, diagnostic tools may underperform for certain ethnic groups if these populations are underrepresented in the training data (Obermeyer et al., 2019). This can lead to misdiagnosis or suboptimal treatment, undermining the quality of care for affected individuals.

Patient Autonomy

The complexity of AI-driven recommendations can make it difficult for patients to understand how decisions are made, potentially undermining informed consent and their ability to participate meaningfully in their own care (London, 2019). When patients are unable to comprehend the rationale behind AI-generated advice, their autonomy is compromised, as they cannot make truly informed choices about their treatment options.

Justice

Biased AI systems may lead to unequal treatment outcomes, challenging the principle of justice in healthcare (Rajkomar et al., 2018). If certain groups consistently receive lower-quality care due to algorithmic bias, existing health inequalities are likely to be reinforced rather than reduced.

Accountability

Determining responsibility for AI-driven decisions is complex, especially when outcomes are unexpected or adverse (London, 2019). The opacity of many AI systems, often described as "black boxes," makes it difficult to attribute accountability when errors occur, raising concerns for both clinicians and patients.

Safeguards

Solutions such as explainable AI, diverse data collection, and transparent communication are recommended to mitigate ethical risks and promote fairness (Rajkomar et al., 2018). Ensuring that

AI systems are trained on representative datasets and that their decision-making processes are understandable to both clinicians and patients is essential for supporting ethical healthcare innovation.

These findings point out the complexity of integrating AI into healthcare ethically. London (2019) points out that while AI holds great promise for improving outcomes, without careful design and oversight, it risks marginalizing vulnerable populations. Similarly, Obermeyer et al. (2019) stress the need for multidisciplinary collaboration to ensure AI technologies uphold core ethical principles such as autonomy and justice.

Podcast Plan

Segment	Content Summary	Ethical Principles Addressed
Introduction	Host introduction, brief overview of AI in healthcare, and statement of the main ethical issue: algorithmic bias	N/A
Background Context	Explanation of AI's role in healthcare, examples of applications, and discussion of its local and global relevance	N/A
Ethical Analysis	In-depth exploration of algorithmic bias, analysis through autonomy and justice, discussion of tensions and trade-offs	Autonomy and Justice
Stakeholders and Impact	Identification of affected groups (patients, providers, policymakers), discussion of potential consequences of inaction	Autonomy, Justice and Accountability

Reflection and Recommendations	Proposal of safeguards (e.g., explainable AI, inclusive data), reflection on the importance of ethics in innovation	Autonomy, Justice and Accountability

The podcast's structured flow is designed to gradually build listener understanding, moving from foundational knowledge to deeper ethical analysis. This approach supports engagement and comprehension, making complex ethical issues accessible, as recommended by Rajkomar et al. (2018), who emphasize the importance of clear communication in health education to foster informed public discourse.

Podcast Structure and Flow

The podcast begins with a concise introduction, establishing the host and the central ethical issue of algorithmic bias in AI healthcare technologies. The background context segment provides listeners with an accessible overview of how AI is currently used in healthcare, with real-world examples to illustrate its growing influence both locally and globally.

The ethical analysis forms the core of the episode, examining algorithmic bias through the lenses of autonomy and justice. This section discusses how these principles can come into tension, especially when technological efficiency may conflict with the need for fairness and patient empowerment.

Next, the stakeholders and impacts segment identifies those most affected by biased AI systems, including patients from marginalised backgrounds, healthcare providers, and policymakers. The discussion highlights the risks of ignoring these issues, such as eroding trust in healthcare and perpetuating health disparities. The episode concludes with a reflection and recommendations, suggesting practical safeguards like explainable AI and representative data collection. The host reflects on the broader importance of ethical considerations in guiding healthcare innovation, emphasising that technological progress should serve the needs and rights of all patients.

Conclusion

Addressing algorithmic bias in AI healthcare technologies is essential to ensure that these innovations benefit all patients fairly and respectfully. The ethical principles of autonomy and justice provide a valuable framework for understanding the challenges and guiding the development of more equitable AI systems. By prioritizing transparency, inclusivity, and accountability, healthcare providers and developers can work together to build trust and improve health outcomes. Ultimately, embedding ethics at the heart of AI innovation will help create a healthcare system that is both technologically advanced and deeply humane.

Reference List

London, A.J. (2019) ‘Artificial intelligence and black-box medical decisions: accuracy versus explainability’, Hastings Center Report, 49(1), pp. 15–21. Available at: https://onlinelibrary.wiley.com/doi/full/10.1002/hast.983

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

Rajkomar, A., Hardt, M., Howell, M.D., Corrado, G. and Chin, M.H. (2018) ‘Ensuring fairness in machine learning to advance health equity’, Annals of Internal Medicine, 169(12), pp. 866–872. Available at: https://www.acpjournals.org/doi/10.7326/M18-1990