The Algorithmic Redlining of Health: How Biased Healthcare AI Perpetuates and Deepens Racial Disparities

Jennifer Adane

06/27/2023

Introduction

Artificial intelligence (AI) is heralded as a revolutionary force in healthcare, promising precision, efficiency, and personalization. Yet behind this promise lies a sobering reality: AI systems are not immune to the social biases embedded in the data from which they learn. In fact, they can amplify them. One emerging concern is *algorithmic redlining*—a modern analogue of discriminatory lending practices—where AI systems in healthcare reproduce and entrench racial disparities under the pretense of neutrality.

How Bias Gets Baked Into the Algorithm

AI models are fundamentally data-driven. In healthcare, these models often rely on historical patient records which are riddled with inequities—reflecting disparities in access, diagnosis, and treatment. For example, Obermeyer et al. (2019) demonstrated that a widely used algorithm, intended to allocate additional care to the sickest patients, prioritized white patients over Black patients due to its reliance on healthcare expenditure as a proxy for health need. Because Black patients historically have less access to healthcare and thus lower spending, they were erroneously assigned lower risk scores.

Similarly, AI tools in dermatology, which are often trained on images of light-skinned individuals, perform poorly on darker skin tones (Adamson and Smith, 2018). This can result in delayed or incorrect diagnoses for skin cancers and other dermatologic conditions in people of color. In the realm of emotion or pain recognition, facial analysis tools have been shown to underperform on non-white faces, leading to disparities in pain management and mental health diagnosis (Buolamwini and Gebru, 2018).

The Mechanism of Harm: Proxy Variables and Historical Injustice

Algorithmic redlining is not just a design flaw—it's a reflection of structural racism. AI models often rely on proxy variables such as zip codes, spending history, or frequency of medical visits, which correlate with race and socioeconomic status. When these are uncritically included in predictive models, they reproduce the systemic disadvantages of marginalized communities.

Additionally, the lack of diversity among developers and in datasets compounds these problems. Technical teams that lack cultural competency may overlook bias, while the exclusion of marginalized voices in development processes reduces the ability to detect and correct such issues.

Real-World Impact: Inequity in Outcomes

These biases are not theoretical—they have tangible consequences. Biased models in emergency departments have led to under-triage of Black patients, resulting in fewer diagnostic tests and interventions (Green et al., 2007). This creates a feedback loop: communities that receive poorer care continue to have worse outcomes, which reinforces their lower prioritization in future models.

Systemic Solutions: Breaking the Cycle of Bias

We must not reject AI in healthcare; rather, we must build it better. First, AI systems should undergo rigorous audits, disclosing training data, decision-making logic, and performance across demographic groups. Second, data must be diversified—beyond representation, context matters. Training sets should reflect varied social determinants of health.

Third, interdisciplinary collaboration is vital. Technologists must work with ethicists, community leaders, and public health experts to ensure models align with principles of justice. Fourth, robust regulatory frameworks must be developed. Presently, many healthcare AI tools evade regulatory oversight unless deemed medical devices. Equity-centered regulation must close this gap.

Conclusion

AI in healthcare holds transformative potential—but only if it is designed with equity at its core. Algorithmic redlining is not merely a side effect of flawed models; it is a systemic problem that demands systemic solutions. By addressing bias proactively, we can build tools that enhance—not erode—justice in health.

The Bibliography

References

Adamson, A. S., & Smith, A. (2018). Machine learning and health care disparities in dermatology. *JAMA Dermatology*, 154(11), 1247–1248. https://doi.org/10.1001/jamadermatol.2018.2348

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 81, 77–91.

Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L. I., & Banaji, M. R. (2007). Implicit bias among physicians and its prediction of thrombolysis decisions for Black and white patients. *Journal of General Internal Medicine*, 22(9), 1231–1238. https://doi.org/10.1007/s11606-007-0258-5

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