



Math in Social Context: Racial Bias in Health Care Machine Learning Algorithms



Esther Akinpelu

What is the Problem?

Machine learning algorithms in healthcare are racially biased, especially when diagnosing patients and choosing candidates for care management programs.

Context

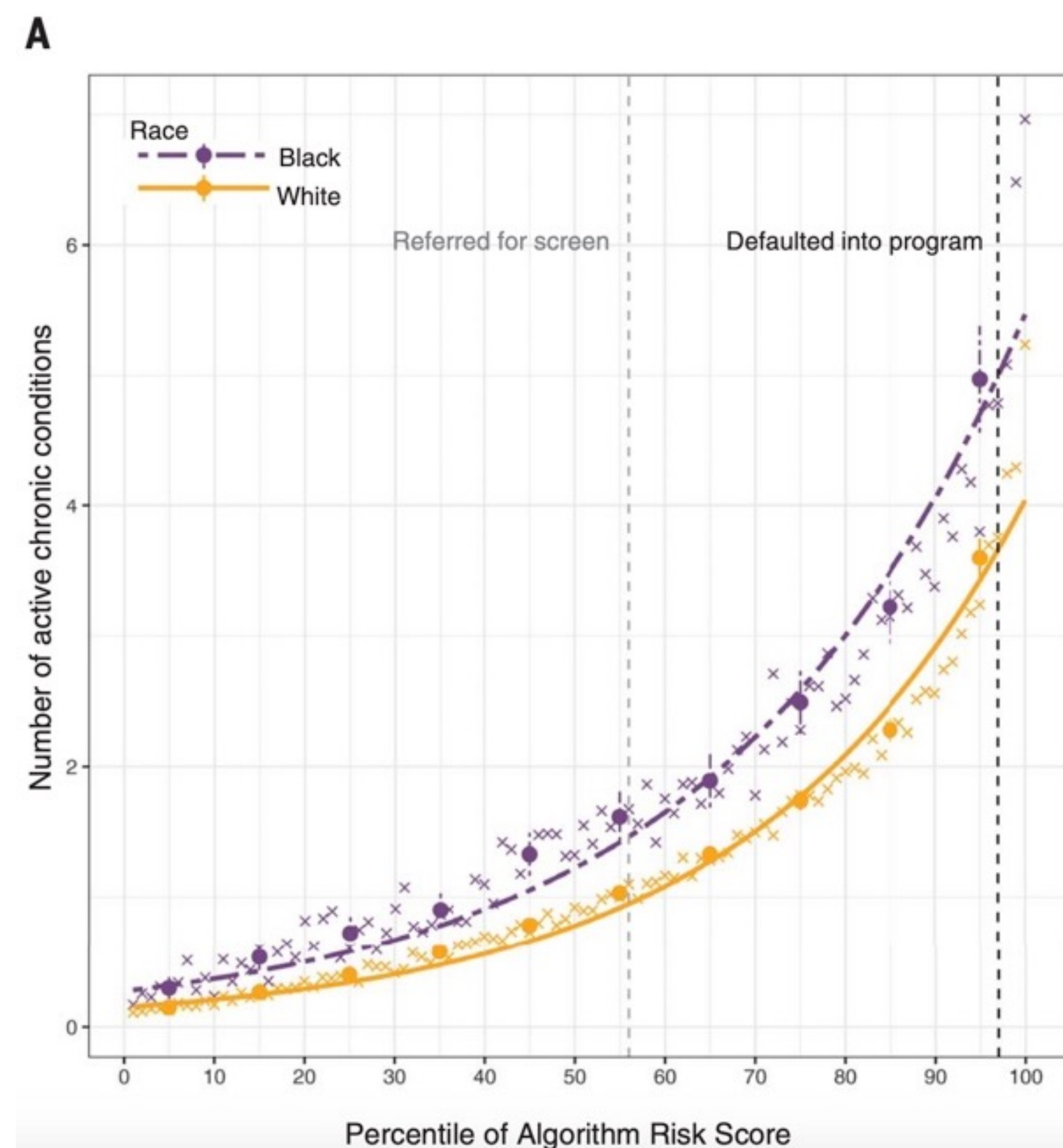
- Machine learning algorithms are tools used for predicting information using different types of data (race, gender, socioeconomic status, etc..)
- Algorithms will find patterns in datasets that could translate into discrimination.
- Algorithms used in healthcare can help in diagnosing diseases.
- Health care management programs are programs that provide better care for patients with high-risk chronic diseases who are most likely to benefit more from the program.
- Decisions about who should enroll in these programs are decided with the help of algorithms that predict high risk scores for patients.
- Usually, these algorithms are created using past data.

Misdiagnoses

- The Kidney Donor Risk Index predicted that African Americans have higher glomerular filtration, which means that they have better kidney functions than Caucasians.
- The Convolutional Neural Network (CNN) is a medical algorithm used to classify skin lesions. It is accurate in classifying skin lesions on lighter skins than darker skins.

Care Management

- Obermeyer et al. study of how well the algorithm is calibrated (does what we measure the same as the predicted risk score) across race
- Overall health status (number of active chronic conditions) of white and black patients in the study compared to their percentile of their algorithm risk score
- The mean number of chronic conditions for blacks were higher than that of whites when compared to the percentile of algorithm risk score.
- More discrepancies between the algorithm's predicted risk factors and the mean number of active chronic illnesses between blacks and whites.
- With more context, patients above the 55th percentile were referred to their primary doctor to make the decision.
- Patients above the 97th percentile were automatically enrolled in program.



Implications

- Misdiagnosis could be life or death.
- Black patients with melanoma have a higher mortality rate than white patients.
- Algorithm risk scores could affect decisions for who gets accepted into the care management programs.
- African Americans with low predicted risk scores but high active chronic illnesses might miss out on enrollment into care management programs.
- Black patients that have high predicted risk scores but low active chronic conditions, might get enrollment into the programs.

How Do We Fix the Racial Bias?

- There needs to be a way to quantify a standard of fairness for algorithms to recognize.
- Transparency in how algorithms work will show medical professionals what factors the algorithms are using to make predictions. This will allow medical professionals to decide which factors are more important in making medical decisions for patients.
- Making sure the input data is unbiased as much as possible (i.e., same number of patients of different races, gender, etc.). Engineers and scientists can re-use and validate algorithms

References

- 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>
- O'Neil, C., & Gunn, H.K. (2020). Near-Term Artificial Intelligence and the Ethical Matrix. *Ethics of Artificial Intelligence*.
- Hoffman, S., & Podgurski, A. (2020). Artificial intelligence and discrimination in health care. *Yale Journal of Health Policy, Law and Ethics*, 19(3), 1-49.
- Norori, N., Hu, Q., Aellen, F. M., Faraci, F. D., & Tzovara, A. (2021). Addressing bias in big data and AI for health care: A call for open science. *Patterns*, 2(10), 100347. <https://doi.org/10.1016/j.patter.2021.100347>