Algorithmic Bias & Racial Disparities in Medicine

Kevin Hoang
Peter Gao (Mentor)

UNIVERSITY *of* WASHINGTON

Racial Biases in Medicine

- History of medical mistreatment (J. Marion Sims, Tuskegee Experiment, Forced Sterilization of Hispanic, Black, and Native American Women)
- Black patients undertreated for pain, underrepresented in drug & vaccine trials, pulse oximeters less accurate for those with darker skin
- Minority groups have worse health outcomes than whites

Usage of Algorithms in Medicine

- Algorithms can range from formulas made by humans to AI/ML computer models
- Medical imaging AI can outperform specialists in cancer detection, osteoarthritis diagnosis, and vision-loss prediction
- AI can predict, diagnose, or prognose diseases with high accuracy

Flaws of Algorithms

Table 1. Examples of Race Correction in Clinical Medicine.*

Tool and Clinical Utility	Input Variables	Use of Race	Equity Concern
Cardiology			
The American Heart Association's Get with the Guidelines–Heart Failure[9] (https://www.mdcalc.com/gwtg-heart-failure-risk-score) *Predicts in-hospital mortality in patients with acute heart failure. Clinicians are advised to use this risk stratification to guide decisions regarding initiating medical therapy.*	Systolic blood pressure Blood urea nitrogen Sodium Age Heart rate History of COPD Race: black or nonblack	Adds 3 points to the risk score if the patient is identified as nonblack. This addition increases the estimated probability of death (higher scores predict higher mortality).	The original study envisioned using this score to "increase the use of recommended medical therapy in high-risk patients and reduce resource utilization in those at low risk."[9] The race correction regards black patients as lower risk and may raise the threshold for using clinical resources for black patients.
Cardiac surgery			
The Society of Thoracic Surgeons Short Term Risk Calculator[10] (http://riskcalc.sts.org/stswebriskcalc/calculate) *Calculates a patient's risks of complications and death with the most common cardiac surgeries. Considers >60 variables, some of which are listed here.*	Operation type Age and sex Race: black/African American, Asian, American Indian/Alaskan Native, Native Hawaiian/Pacific Islander, or "Hispanic, Latino or Spanish ethnicity"; white race is the default setting.	The risk score for operative mortality and major complications increases (in some cases, by 20%) if a patient is identified as black. Identification as another nonwhite race or ethnicity does not increase the risk score for death, but it does change the risk score for major complications such as renal failure, stroke, and prolonged ventilation.	When used preoperatively to assess a patient's risk, these calculations could steer minority patients, deemed higher risk, away from these procedures.

Vyas et al. (2020)

- **Important to understand algorithms aren't perfect**
- **Bad training data or flawed input variables**
- **Unexplainable conclusions from "black box" models**

Case Study: Risk Score Algorithms in Medicine

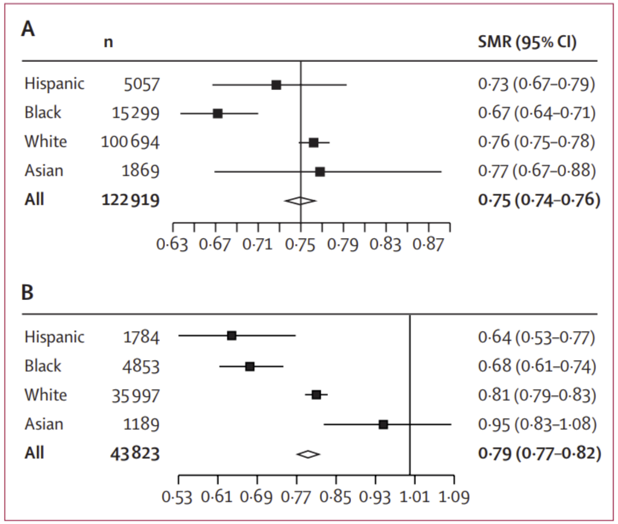
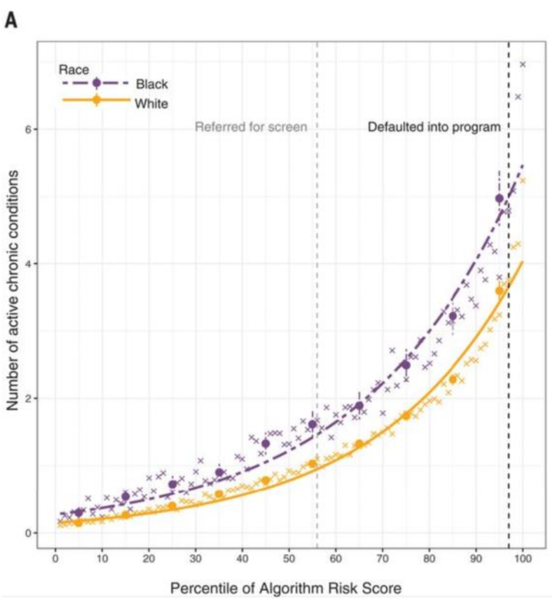
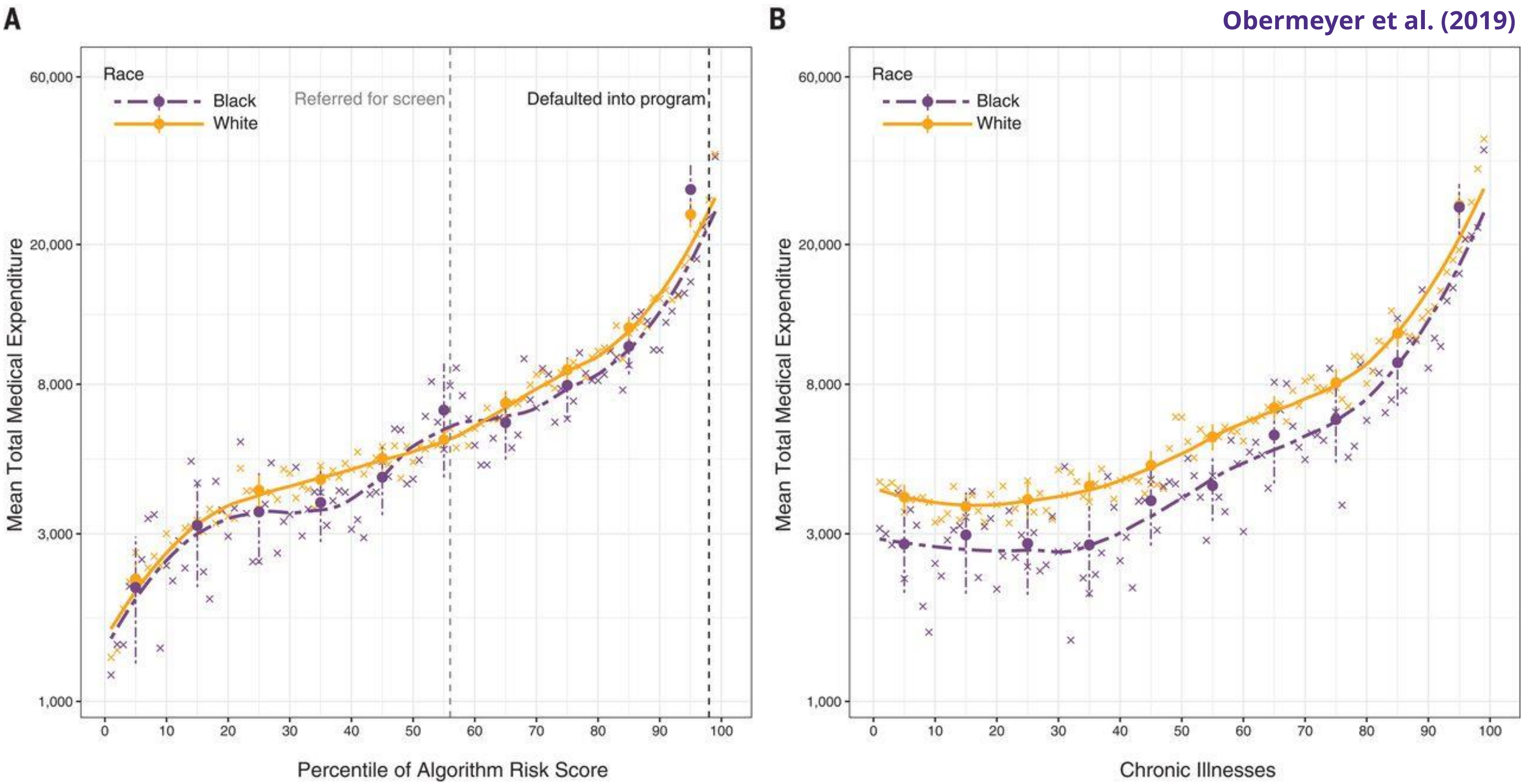


Figure 3: SMR for APACHE score in the eICU-CRD and OASIS score in MIMIC-III across ethnicities



- Sarkar et al. (2021), "Performance of intensive care unit…"
  - ➢ Severity scoring systems for resource allocation (COVID-19)
  - ➢ Hispanic & Black mortality overpredicted, less care allocated
- Obermeyer et al. (2019), "Dissecting Racial Bias…"
  - ➢ Commercial "high- risk care management" algorithm based on medical expenditure history

Obermeyer et al. (2019)



- **No racial disparities when comparing medical expenditure to risk score**
- **However, medical expenditure is a bad metric for health**

Mathematical Definitions of Fairness

- Equal Patient Outcomes
- Equal Performance
  - Equal Accuracy
  - Equal False Positive
  - Equal False Negative
- Equal Allocation

Transparency & Interpretability

- Transparency – Providing training dataset and code
- Transparent algorithms can be "auditable"
- Interpretable algorithms' decisions can be explained
- Interpretability and Performance tradeoff?

Citations

- Vyas, D. A., Eisenstein, L. G., & Jones, D. S. (2020). Hidden in plain sight—reconsidering the use of race correction in clinical algorithms.
- Sarkar, R., Martin, C., Mattie, H., Gichoya, J. W., Stone, D. J., & Celi, L. A. (2021). Performance of intensive care unit severity scoring systems across different ethnicities in the USA: a retrospective observational study. The Lancet Digital Health, 3(4), e241-e249.
- 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.

Acknowledgements

- Peter Gao, my DRP mentor
- UW Statistics and Probability Association
- Directed Reading Project Mentees

Questions or Comments?