Explaining Treatment Disparities from a Causal Perspective with EHRs

Linying Zhang, MS[1], Xinzhuo Jiang, MS[1], Karthik Natarajan, PhD[1,2], George Hripcsak, MD, MS[1,2]

[1]Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY; [2]Medical Informatics Services, NewYork-Presbyterian Hospital, New York, NY.

Introduction

Ensuring fairness in treatment allocation is a crucial step towards achieving equitable healthcare. The decision-making process for treatment is complex and involves several factors, including a patient's medical history, environmental and social factors, and, potentially, race and sex. In health disparities research, women and racial and ethnic minorities are often found to receive less treatment in various clinical domains, raising concerns about inequity in clinical practice. However, research on health disparities often relies on correlational analysis, which may not reveal the underlying causal mechanisms behind disparities. The critical question is: How can we explain observed disparities found in the data in terms of the unobservable causal mechanisms?

In this study, we explored a causal fairness analysis framework [1] to better understand the causes and their impact on treatment disparities. As an example, we focused on assessing treatment disparities for coronary artery disease (CAD), as previous studies have shown that women, racial and ethnic minorities, patients without health insurance, and those living in low-income neighborhoods are associated with lower use of treatment for CAD. We quantified the effect of race on treatment allocation by decomposing the effect into direct effect, indirect effect (via baseline medical conditions with or without social determinants of health (SDoH)), and confounded effect (via age and sex) using the causal fairness framework.

Methods

Data, Cohort, and Features Data used for the analysis come from Columbia University Irving Medical Center (CUIMC) electronic health record (EHR) database v Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) v5.4. The coronary artery disease (CAD) cohort was phenotyped based the presence of coronary arteriosclerosis diagnosis codes. Patients with missing race or gender were excluded from the study. The treatment group was defined as patients treated with either percutaneous coronary intervention (PCI) or coronary artery bypass grafting (CABG). The index date was the date of the initial treatment in a patient's EHR. The rest of the patients were in the control group. The index date for the control group was the latest clinical visit with a coronary arteriosclerosis diagnosis. We extracted demographics (race, sex, age), medical conditions, and insurance plans within one-year prior to index date. We also extracted area deprivation index (ADI) by linking geocoded addresses in the EHR with Census data.

Fairness Measures and Algorithms We assume the standard fairness model as in Fig.1(a). Based on the causal mechanisms, the total effect of race on treatment allocation can be decomposed into direct effect (Fig.1b), indirect effect (Fig.1c), and confounded effect (Fig.1d). These effects can be expressed mathematically as follows.

$$Total = E[Y|X = x_1] - E[Y|X = x_0], \qquad Direct = E[Y_{x_1,W_{x_0}} \mid x_0] - E[Y_{x_0} \mid x_0]$$

$$Indirect = E[Y_{x_1,W_{x_0}} \mid x_0] - E[Y_{x_1} \mid x_0], \qquad Confounded = E[Y_{x_0} \mid x_1] - E[Y_{x_0} \mid x_0]$$

$$Total = Direct - Indrect - Confounded$$

For estimating the causal measures of fairness, we implemented random forest to estimate causal quantities in the forms of $E[Y_x]$ or $E[Y_{x_1,W_{x_0}}]$. The estimator is doubly-robust, meaning that the estimator will be consistent as long as one model involved in the estimator is consistent, a highly desired property for estimators.

Results

The final cohort consists of 41,630 patients, including 8,428 (20.2%) patients in the treatment group and 33,202 (79.8%) patients in the control group. In the treatment group, 6,437 patients had PCI treatment and 1,991 patients had CABG treatment. There were 6298 patients (15.2%) whose race variable was recorded as Black or African American.

Fig.2 presents the decomposition of racial disparity in revascularization based on causal mechanisms, which help us understand the different components contributing to the observed treatment disparities between Black and non-Black patients. The total disparity indicates that Black patients were 5.0% [95%CI: 4.1% to 6.0%] less likely to receive revascularization treatment compared to non-Black patients. This total variation can be broken down into three causal components: direct effect, indirect effect, and confounded effect.

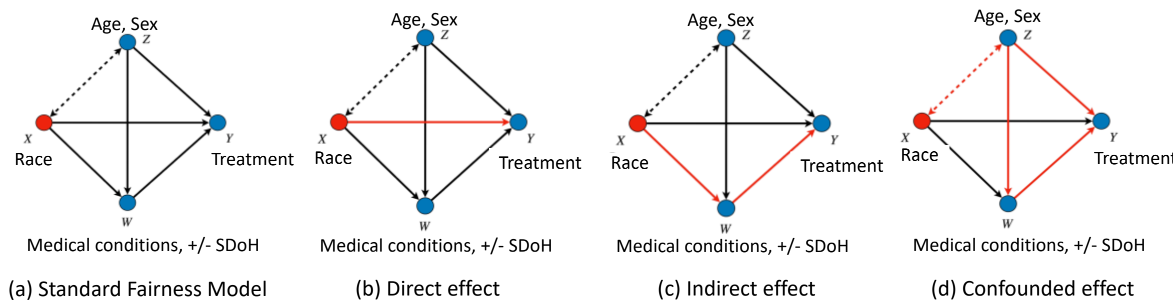

Figure 1. Causal fairness analysis of treatment disparity. (a) Standard fairness model that decomposes racial disparity in treatment allocation based on causal mechanisms. (b) Direct effect of race on treatment allocation. (c) Indirect effect of race on treatment allocation. If SDoH are included in the W-set, then the indirect effect captures both the effect via medical conditions and via SDoH. Otherwise, it captures only the effect mediated by medical conditions. (d) Confounded effect of race on treatment allocation through historical context.

The indirect effect captures the influence of race on treatment allocation mediated by medical conditions, and in som cases, also by SDoH. In this study, when both medical conditions and SDoH were included in the mediator set (Fig.2a), the indirect effect was 2.2% [95%CI: 1.0% to 3.5%], which indicates that if a non-Black patient were Black, their probability of receiving treatment would have decreased by 2.2% due to the changes in their medical conditions and SDoH caused by their race. When SDoH were not included (Fig.2b), this indirect effect was 0.9% [95%CI: -0.6% to 2.2%], 1.3% lower than the one with SDoH, which indicates that SDoH (in this case, ADI and insurance) had a profound impact on the treatment allocation.

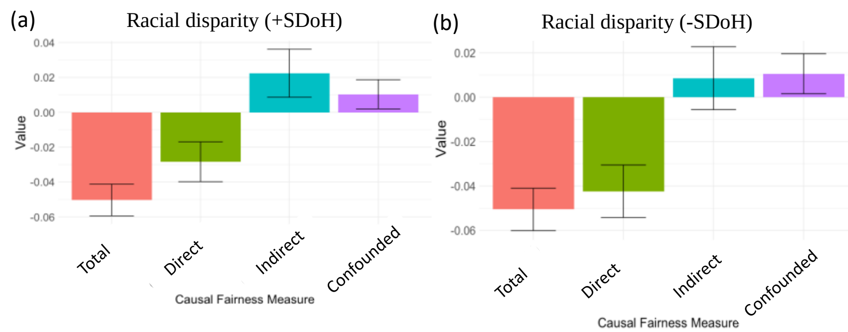

Figure 2. Treatment disparity decomposition for CAD cohort. (a) The decomposition of the effect of race on the disparity of cardiac revascularization *with* SDoH. The indirect effect reflects the effect of race mediated by both medical history and SDoH. (b) The decomposition of the effect of race on the disparity of cardiac revascularization *without* SDoH. The indirect effect reflects the effect of race mediated by medical history alone.

The confounded effect accounts for the impact of other demographic factors (age, sex) on treatment allocation. Based on Fig.2a, the confounded effect was 1.0% [95%CI: 0.1% to 2%], suggesting that other demographic factors contributed to a 1.0% difference in probability of receiving treatment between Black and non-Black patients. This effect measures the role of other demographic factors in explaining the observed disparities in treatment allocation.

The direct effect measures any residual impact of race on treatment allocation that is not accounted by the other demographic factors (Z-set), medical conditions, and SDoH (W-set). In the case when SDoH were included (Fig.2a), the direct effect was 3.0% [95%CI: 2.0% to 4.0%], suggesting that if a non-Black patient were Black, their probability of receiving treatment would have decreased by 3.0%, assuming all other features remain the same. Note that race does have a statistically significant direct effect on the treatment allocation.

Conclusions Our study demonstrates the utility of a causal fairness analysis framework in identifying and quantifying the impact of race on treatment allocation in clinical practice. By decomposing the impact of race into direct, indirect, and confounded effects, we were able to explain the observed treatment disparities based on various causal mechanisms. We advocate for future research to focus on the integration of more comprehensive SDoH data into the causal analysis of treatment disparities, further refining our understanding of health disparities towards building an equitable healthcare.

References

[1] Plecko, D., & Bareinboim, E. (2022). Causal fairness analysis. arXiv preprint arXiv:2207.11385.