

RESEARCH

Open Access



Propensity score analysis for health care disparities: a deweighting approach

Byeong Yeob Choi^{1*}

Abstract

Background Propensity score weighting is a useful tool to make causal or unconfounded comparisons between groups. According to the definition by the Institute of Medicine (IOM), estimates of health care disparities should be adjusted for health-status factors but not for socioeconomic status (SES) variables. There have been attempts to use propensity score weighting to generate estimates that are concordant with IOM's definition. However, the existing propensity score methods do not preserve SES distributions in minority and majority groups unless SES variables are independent of health status variables.

Methods The present study introduces a deweighting method that uses two types of propensity scores. One is a function of all covariates of health status and SES variables and is used to weight study subjects to adjust for them. The other is a function of only the SES variables and is used to deweight the subjects to preserve the original SES distributions.

Results The procedure of deweighting is illustrated using a dataset from a right heart catheterization (RHC) study, where it was used to examine whether there was a disparity between black and white patients in receiving RHC. The empirical example provided promising evidence that the deweighting method successfully preserved the marginal SES distributions for both racial groups but balanced the conditional distributions of health status given SES.

Conclusions Deweighting is a promising tool for implementing the IOM-definition of health care disparities. The method is expected to be broadly applied to quantitative research on health care disparities.

Keywords Absolute standardized mean difference, Balancing weight, Deweighting, Health care disparity, Health status, Propensity score, Socioeconomic status

Introduction

Evaluating racial disparities needs specific statistical approaches to quantify and find solutions to reduce them. For example, Jackson and VanderWeele [1] investigated how disparities in wages would change if disparities in education were removed while preserving the association of childhood SES with race. This approach allows us

to estimate how well removing disparities in education, but not childhood SES, reduces disparities in adulthood wages. Ben-Michale et al. [2] evaluated differences between black and white patients in common outcomes after emergency general surgery (EGS) treatment using an administrative database of hospital claims. They used linear weighting estimators that re-weight white patients to have a similar distribution of the adjustment set to black patients. They found that adjusting for hospitals in addition to patient characteristics substantially reduced the disparity estimate resulting from adjusting for basic demographics (age and sex). This finding implied that hospital-specific factors are important sources of racial

*Correspondence:

Byeong Yeob Choi
choib@uthscsa.edu

¹ Department of Population Health Sciences, UT Health San Antonio, 7703 Floyd Curl Drive, Mail Code 7933, San Antonio 78229, TX, USA



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

disparities in EGS, and thus, interventions targeted at hospital quality may be critical.

In this study, we utilize the definition of the Institute of Medicine (IOM) to propose new disparity estimators. The Institute of Medicine (IOM) defines a health care disparity as the difference in the quality of health care provided to members of racial or ethnic minorities that is not due to their health status, clinical needs, or treatment preferences [3]. To implement the IOM definition, a group of researchers proposed a specific statistical approach that adjusts for health-status factors but not for nonhealth variables, i.e., socioeconomic status (SES) variables, such as education, income, and geographic region [4–6]. The rationale of this approach is that the IOM definition of disparities should include all racial differences in use mediated through SES-related factors. However, we acknowledge that this approach, adjusting for only health-status variables, is one of the ways to implement the IOM definition. Considering such a limitation and the complexity of disentangling the effects of race/ethnicity from SES factors, we have compared the proposed approach with those adjusting for all covariates and investigated how these methods give different results in a real data example.

Several researchers devised statistical models based on PS models to address the needs required by the IOM definition of racial disparities. Cook et al. [5] defined distributional properties necessary for a disparity method to be concordant with the IOM definition of racial disparities in health care use. Specifically, they compared a counterfactual white population with black health status and white SES with the factual black population. This white distribution is counterfactual because it has black health status but white SES. One of their disparity methods to construct the hypothetical white distribution is the health-status PS method. In racial disparity studies, minority race is the “treatment” of interest for the PS. The health-status PS uses only the health-status variables to predict the probability of minority race. Motivated by a disparity study of Li et al. [7], Choi et al. [8] extended the health-status PS method by adopting the balancing weights approach [9] to address the broader range of disparity measures. However, the fundamental limitation of the health-status PS method is that SES variables are not preserved unless they are independent of health status.

In this study, we propose weights that can realize IOM-concordant joint distributions of health status and SES. The key idea is that we weight the groups with the PSs using all covariates but deweight them with the PSs based on only SES. In this way, we directly undo the undesired weighting of the SES variables. Considering this feature of

our approach, we call it a deweighting approach. Importantly, we formally define the target estimands for health disparities and corresponding consistent estimators based on the proposed weights.

The deweighting approach is illustrated in a study of right heart catheterization (RHC) [10] to investigate whether there was a disparity between black and white patients, admitted to the intensive care unit, in receiving RHC. Based on unadjusted proportions, black patients were 1.86% less likely to receive RHC than white patients. However, this unadjusted difference does not account for differences in health status or clinical needs, which do not contribute to disparities. In the RHC study, white patients tended to have lower blood pressure than black patients, and this could contribute to apparent disparities for black patients in receiving RHC. Because the disparities due to clinical needs such as low blood pressure are allowable, these factors need to be adjusted for. Traditional multivariable methods adjust for all available covariates, which also removes differences due to non-clinical needs. Thus, the controlled differences will better represent racial disparities if the SES variables are preserved in both racial groups so that the SES variables can mediate the relationship between race and RHC treatment. This motivates the development of our deweighting approach.

Methods

Notation

We introduce notation to describe our deweighting approach. For unit $i = 1, \dots, N$, let Z_i be the binary variable for minority race/ethnicity and Y_i be the observed health care outcome. In our numerical study, we assessed whether there was a disparity in receiving RHC at the time of admission to the intensive care unit between black and white patients: $Y_i = 1$ if the patient received RHC, and $Y_i = 0$ otherwise. Therefore, in our analysis, $Z_i = 1$ and $Z_i = 0$ indicate black and white patients, respectively. For causal inference, $Y(1)$ and $Y(0)$ are defined as the potential outcomes corresponding to treatment levels 1 and 0. The observed outcome is then $Y = ZY(1) + (1 - Z)Y(0)$ under the consistency assumption.

Let X_i be a row vector of $J + K$ covariates consisting of SES and health-status variables: $X_i = (X_i^s, X_i^h)$, where $X_i^s = (X_{i1}^s, \dots, X_{ij}^s)$ is a row vector of J SES variables, and $X_i^h = (X_{i,J+1}^h, \dots, X_{i,J+K}^h)$ is a row vector of K health-status variables. Let $f(x)$ and $f_z(x)$ denote the densities of X in the whole population and subpopulation of $Z = z$, and

$f(x^s)$ and $f_z(x^s)$ denote those of X^s in the whole population and subpopulation of $Z = z$, respectively, where $z \in \{0, 1\}$. We also define $f(x^h | x^s)$ and $f_z(x^h | x^s)$ to be the conditional densities of X^h given $X^s = x^s$ for the whole population and subpopulation of $Z = z$.

Propensity scores

We define the full PS as a function of all available covariates:

$$e(x) = \Pr(Z = 1 | X = x).$$

In health disparities research, we are interested in comparing health care outcomes between the minority group (coded as $Z = 1$) with the majority group (coded as $Z = 0$). Therefore, in our study, the PS represents the probability of being in the minority group conditional on the covariates. The standard causal inference approach adopts the full PS to balance all covariates between comparison groups. However, balancing all covariates does not comply with the IOM’s definition because the original differences in the SES variables cannot be preserved. To overcome this issue from using the full PS, Cook et al. [5] proposed using the health-status PS, which is a function of only the health-status variables:

$$e(x^h) = \Pr(Z = 1 | X^h = x^h).$$

Choi et al. [8] extended the approach of Cook et al. [5] to study various estimands of health care disparities. Weighting based on the health-status PS is optimal if the health status and SES variables are independent of each other, but it generally alters the SES variables, as demonstrated by Choi et al. [8]. To address this issue, Choi et al. [8] introduced a data-adaptive method to find the balancing weights that minimize the alternations of the SES variables while preventing severe imbalance in the health-status variables. This data-adaptive method might be effective in overcoming the inherent limitation of the health-status PS method but is still an indirect way to preserve SES variables.

Rather than using the health-status PSs, the proposed deweighting approach weights the subjects with the full PSs but deweights them using the PSs based on only the SES variables. Because the SES variables are involved in the weights, the proposed approach directly preserves the SES variables, while adjusting for the health-status variables. We will refer to the PS using only the SES variables as the SES-PS, defined as:

$$e(x^s) = \Pr(Z = 1 | X^s = x^s).$$

Balancing weights

To use the PSs to study health disparities, we adopt the balancing weights approach [9]. The balancing weights based on the full PSs are defined as

$$\omega_1^B(x) = \frac{g(x)}{e(x)}, \text{ and } \omega_0^B(x) = \frac{g(x)}{1-e(x)} \tag{1}$$

where $g(x)$ is a selection function that determines the target population of interest. If we weight the treated group by $\omega_1^B(x)$ and the control group by $\omega_0^B(x)$, then $f_1(x)$ and $f_0(x)$ will be balanced toward $t(x) = g(x)f(x)/E\{g(x)\}$. Typically, $g(x)$ is a function of $e(x)$. Therefore, it is worthwhile to look at the target population when $g(x)$ is equal to $e(x)$ or $1 - e(x)$:

$$f_1(x) \propto e(x)f(x), \text{ and } f_0(x) \propto \{1 - e(x)\}f(x) \tag{2}$$

Based on Eq. (2), we can demonstrate that $f_1(x)\omega_1^B(x) \propto g(x)f(x)$ and $f_0(x)\omega_0^B(x) \propto g(x)f(x)$, and therefore $\omega_1^B(x)$ and $\omega_0^B(x)$ are called the balancing weights.

The specification of $g(x)$ determines the target populations and estimands, as shown in Table 1 of Li et al. [9]. For example, when $g(x) = 1$, the estimand is the average treatment effect (ATE), and the weights are the inverse probability weights (IPWs), $\{\omega_1^B(x), \omega_0^B(x)\} = \{1/e(x), 1/(1 - e(x))\}$. When $g(x) = e(x)$, the estimand is the average treatment effect on the treated (ATT), and the weights are the standardized mortality ratio weights (SMRWs), $\{\omega_1^B(x), \omega_0^B(x)\} = \{1, e(x)/(1 - e(x))\}$. Li et al. [9] introduced overlap weights to identify the average treatment effect in the overlap population (ATO), where $g(x) = e(x)\{1 - e(x)\}$, and the weights are $\{\omega_1^B(x), \omega_0^B(x)\} = \{1 - e(x), e(x)\}$. The overlap weights are useful to address extreme PSs [11].

Let $\tau(x) = E[Y | Z = 1, X = x] - E[Y | Z = 0, X = x]$ denote the conditional average controlled difference given $X = x$ [12]. Based on Hirano et al. [13] and Li et al. [9], we can express the target estimand as the weighted average of $\tau(x)$ over the target population, $g(x)f(x)$:

$$\tau_g = \frac{\int \tau(x)g(x)f(x)dx}{\int g(x)f(x)dx} \tag{3}$$

In causal inference, τ_g is called the weighted average treatment effect (WATE). The necessary conditions to identify the WATE from the observed data include the assumptions of unconfoundedness and positivity [14]. For descriptive comparisons, we will call it the weighted

Table 1 Baseline characteristics of the overall, black non-RHC, black RHC, white non-RHC, and white RHC patients

	Overall	Black non-RHC	Black RHC	White non-RHC	White RHC
N	5380	585	335	2753	1707
Age (mean (SD))	61.99 (16.46)	57.42 (17.79)	54.88 (16.87)	63.45 (16.64)	62.60 (14.90)
Sex= Male (%)	2994 (55.7)	290 (49.6)	165 (49.3)	1505 (54.7)	1034 (60.6)
Cancer (%)					
Metastatic	362 (6.7)	49 (8.4)	18 (5.4)	198 (7.2)	97 (5.7)
No cancer	4105 (76.3)	466 (79.7)	276 (82.4)	2029 (73.7)	1334 (78.1)
Localized	913 (17.0)	70 (12.0)	41 (12.2)	526 (19.1)	276 (16.2)
Disease category (%)					
ARF	2354 (43.8)	255 (43.6)	121 (36.1)	1244 (45.2)	734 (43.0)
CHF	422 (7.8)	42 (7.2)	37 (11.0)	193 (7.0)	150 (8.8)
MOSF	1504 (28.0)	146 (25.0)	140 (41.8)	557 (20.2)	661 (38.7)
Other	1100 (20.4)	142 (24.3)	37 (11.0)	759 (27.6)	162 (9.5)
Number of comorbidities (mean (SD))	1.51 (1.16)	1.42 (1.17)	1.35 (1.07)	1.56 (1.17)	1.50 (1.14)
Duke activity status index (mean (SD))	20.48 (5.31)	20.45 (5.81)	20.82 (5.25)	20.30 (5.39)	20.73 (5.00)
APACHE score (mean (SD))	54.59 (19.88)	52.56 (19.99)	62.82 (20.94)	50.48 (18.30)	60.31 (20.21)
Glasgow coma score (mean (SD))	21.03 (30.48)	26.06 (32.47)	21.63 (30.65)	21.50 (31.26)	18.44 (28.14)
Mean blood pressure (mean (SD))	78.56 (38.16)	93.76 (42.08)	70.64 (39.42)	83.10 (38.00)	67.59 (33.28)
WBC (mean (SD))	15.77 (11.98)	15.87 (13.25)	16.54 (14.36)	15.27 (11.09)	16.40 (12.36)
Heart rate (mean (SD))	114.98 (41.52)	113.88 (42.04)	118.64 (46.20)	112.58 (40.93)	118.52 (41.06)
Respiratory rate (mean (SD))	28.05 (14.12)	29.86 (15.16)	26.63 (15.29)	28.74 (13.74)	26.61 (13.98)
Temperature (mean (SD))	37.61 (1.78)	37.53 (1.92)	37.34 (2.01)	37.65 (1.71)	37.64 (1.79)
PaO2/FiO2 ratio (mean (SD))	221.31 (114.15)	261.19 (121.04)	214.19 (118.05)	234.86 (114.28)	187.20 (101.30)
Albumin (mean (SD))	3.09 (0.79)	3.13 (0.63)	2.88 (0.76)	3.17 (0.68)	2.99 (0.97)
Hematocrit (mean (SD))	31.95 (8.38)	31.06 (8.72)	29.52 (7.71)	33.19 (8.75)	30.75 (7.40)
Bilirubin (mean (SD))	2.11 (4.37)	2.04 (4.54)	2.25 (3.95)	1.82 (3.81)	2.58 (5.12)
Creatinine (mean (SD))	2.13 (2.05)	2.49 (2.82)	2.87 (2.51)	1.81 (1.80)	2.38 (1.92)
Sodium (mean (SD))	136.81 (7.63)	138.21 (7.69)	136.42 (7.69)	136.81 (7.59)	136.42 (7.60)
Potassium (mean (SD))	4.07 (1.02)	4.11 (1.14)	4.13 (1.09)	4.07 (1.00)	4.03 (0.99)
PaCO2 (mean (SD))	38.80 (13.15)	38.30 (13.24)	36.77 (10.93)	40.40 (14.34)	36.78 (11.00)
PH (mean (SD))	7.39 (0.11)	7.39 (0.10)	7.37 (0.11)	7.39 (0.11)	7.38 (0.11)
Weight (mean (SD))	68.25 (29.05)	65.43 (30.83)	75.01 (27.10)	65.46 (29.17)	72.40 (27.89)
'Do not resuscitate' status on day 1 =Yes (%)	626 (11.6)	63 (10.8)	16 (4.8)	414 (15.0)	133 (7.8)
Insurance (%)					
Government	2338 (43.5)	361 (61.7)	186 (55.5)	1196 (43.4)	595 (34.9)
Private	2752 (51.2)	174 (29.7)	111 (33.1)	1440 (52.3)	1027 (60.2)
Uninsured	290 (5.4)	50 (8.5)	38 (11.3)	117 (4.2)	85 (5.0)
High school education or more (%)	1527 (28.4)	102 (17.4)	69 (20.6)	779 (28.3)	577 (33.8)
Income >=\$25k (%)	1258 (23.4)	63 (10.8)	44 (13.1)	649 (23.6)	502 (29.4)

average controlled difference (WACD). Because race/ethnicity is not manipulable, the estimand more related to health care disparity is the WACD. However, we will adopt the terminologies used for the WATE. For example, we will use the terms 'ATE' and 'ATT' to represent the measures of health care disparity in the combined (minority plus majority) and minority groups, respectively.

Following Theorem 1 of Li et al., the consistent estimator for τ_g is

$$\hat{\tau}_g = \frac{\sum_i \omega_1^B(X_i) Z_i Y_i}{\sum_i \omega_1^B(X_i) Z_i} - \frac{\sum_i \omega_0^B(X_i) (1 - Z_i) Y_i}{\sum_i \omega_0^B(X_i) (1 - Z_i)}$$

Our proposed approach is based on the balancing weights in Eq. (1) but needs additional balancing weights based on the SES-PSs:

$$\omega_1^B(x^s) = \frac{g(x^s)}{e(x^s)}, \text{ and } \omega_0^B(x^s) = \frac{g(x^s)}{1 - e(x^s)}.$$

The key contribution of this study is to demonstrate that the following weights $\omega_1(x) = \frac{\omega_1^B(x)}{\omega_0^B(x)}$ and $\omega_0(x) = \frac{\omega_0^B(x)}{\omega_1^B(x)}$, can create joint distributions aligned with the IOM’s definition. Before we discuss these proposed weights, we will define the estimands for health care disparities.

Estimands for health care disparities

It is important to note that the WATE or WACD in (3) is obtained by averaging $\tau(x)$ over the common target joint distribution of health status and SES. Accordingly, the estimand in (3) may not be appropriate as a measure of health care disparities because the SES variables are altered. Because the IOM definition requires SES to be preserved but health status to be balanced between groups, it would be desirable to balance health status conditional on SES. Therefore, our target population for each group is represented by the target conditional health-status distribution given SES times its own marginal SES distribution.

Let $f_g(x^h | x^s)$ denote a target conditional distribution of health status given SES. We will show that $f_g(x^h | x^s) = f(x^h | x^s)g(x)/g(x^s)$, where $g(x)$ is a selection function used to define the WATE in (3). The selection function $g(x^s)$ is a function of $e(x^s)$ but with the same form as $g(x)$. A simple example arises when $g(x) = g(x^s) = 1$; the target conditional distribution of health status is $f(x^h | x^s)$, which is for the combined population. In this case, the target conditional distribution is the combined distribution of health status from minority and majority subjects, and thus, we will refer to the corresponding estimand as the ATE measure of health disparity. When $g(x) = e(x)$ and $g(x^s) = e(x^s)$, the target conditional distribution of health status is $f_1(x^h | x^s)$, which is for the minority population. Because the target conditional distribution is for minority health status, we will refer to the corresponding estimand as the ATT measure of health disparity. From the fact that a joint distribution is a conditional distribution of health status given SES times a marginal distribution of SES, the target joint distribution for the group $Z = z$ can be written as $t_z(x) = f_g(x^h | x^s)f_z(x^s)/E[f_g(x^h | x^s)f_z(x^s)]$, where $E[f_g(x^h | x^s)f_z(x^s)]$ are normalizing constants for $z = 0,1$. These joint distributions have the common conditional distribution of health status $f_g(x^h | x^s)$ but their own original marginal distributions $f_1(x^s)$ and $f_0(x^s)$ for minority and majority groups, respectively.

Now we define the estimand for disparity as $\tau_g = \mu_{g1} - \mu_{g0}$, where μ_{g1} and μ_{g0} are:

$$\mu_{g1} = \frac{\int E[Y|Z=1, X=x]f_g(x^h|x^s)f_1(x^s)dx}{\int f_g(x^h|x^s)f_1(x^s)dx} \tag{4}$$

and

$$\mu_{g0} = \frac{\int E[Y|Z=0, X=x]f_g(x^h|x^s)f_0(x^s)dx}{\int f_g(x^h|x^s)f_0(x^s)dx}. \tag{5}$$

Note that μ_{g1} and μ_{g0} are defined by averaging the conditional expectation of Y over the same target conditional distribution of health status but different marginal distributions of SES. Our hypothesis of interest is that this τ_g is equal to 0 or not.

In the next section, we will discuss important examples of the proposed weights to demonstrate how μ_{g1} and μ_{g0} can be estimated for the ATE and ATT measures of health disparities, aligned with IOM’s definition.

Proposed weights

We illustrate our approach with two important examples, the ATE and ATT measures of health disparity. Suppose that we used the IPWs, $\{1/e(x), 1/(1 - e(x))\}$, to estimate the ATE. Because the full PS includes SES variables, they could be altered by the weighting. To alleviate the impacts of the IPWs on the SES variables, we deweight the subjects based on the SES-PSs. That is, we multiply the SES-PSs to the minority group and the ‘1 minus SES-PSs’ to the majority group. As a result, we have the following weights:

$$\omega_1(x) = \frac{e(x^s)}{e(x)}, \text{ and } \omega_0(x) = \frac{1 - e(x^s)}{1 - e(x)}. \tag{6}$$

First, we demonstrate that the joint distribution resulting from weighting the minority group with $\omega_1(x)$ is equal to the conditional distribution of health status given SES in the combined group times the minority distribution of SES:

$$f_1(x)\omega_1(x) \propto f(x)e(x^s) = f(x^h|x^s)f(x^s)e(x^s) \propto f(x^h|x^s)f_1(x^s).$$

This resultant joint distribution, $f(x^h|x^s)f_1(x^s)$, is counterfactual because the conditional health-status distribution is from the combined population, but the marginal SES distribution is from the minority group.

In the same way, we can show that the joint distribution resulting from weighting the majority group with $\omega_0(x)$ is equal to the conditional distribution of health status given SES in the combined group times the majority distribution of SES:

$$f_0(x)\omega_0(x) \propto f(x)\{1 - e(x^s)\} = f(x^h|x^s)f(x^s)\{1 - e(x^s)\} \propto f(x^h|x^s)f_0(x^s).$$

Therefore, the proposed weights in Eq. (6) generate hypothetical distributions aligned with the IOM’s definition because the marginal SES distributions for both groups are kept but the target conditional distribution of health status coincides with $f(x^h|x^s)$. Therefore, the original SES differences between the groups can contribute to the ATE estimate of health disparity, while adjusting for health status.

Now we discuss the racial disparity in the minority group, the ATT measure of health disparity. Suppose that

$$f_0(x)\omega_0(x) = f_0(x) \frac{g(x)}{1 - e(x)} \frac{1 - e(x^s)}{g(x^s)} \propto f(x)\{1 - e(x^s)\} \frac{g(x)}{g(x^s)} \propto \left\{ f(x^h|x^s) \frac{g(x)}{g(x^s)} \right\} f_0(x^s).$$

we used the SMRWs, $\{1, e(x)/(1 - e(x))\}$, to estimate the ATT. In this case, we do not deweight the minority group because all the weights are 1, which does not alter the SES variables. Because the majority subjects are weighted by $e(x)/(1 - e(x))$, we multiply the majority group by the inverse of $\frac{e(x^s)}{[1 - e(x^s)]}$ to deweight them. As a result, we have the following weights:

$$\omega_1(x) = 1, \text{ and } \omega_0(x) = \frac{e(x)}{1 - e(x)} \left\{ \frac{e(x^s)}{1 - e(x^s)} \right\}^{-1}. \tag{7}$$

$$f_g(x^h|x^s) = f(x^h|x^s) \frac{e(x)}{e(x^s)} = f(x)e(x)\{f(x^s)e(x^s)\}^{-1} \propto f_1(x)f_1(x^s)^{-1} = f_1(x^h|x^s),$$

We can show that the joint distribution resulting from weighting the majority group with $\omega_0(x)$ in Eq. (7) is equal to the conditional distribution of health status given SES in the minority group times the majority distribution of SES:

$$f_0(x)\omega_0(x) \propto f_1(x) \left\{ \frac{e(x^s)}{1 - e(x^s)} \right\}^{-1} = f_1(x^h|x^s)f_1(x^s) \frac{1 - e(x^s)}{e(x^s)} \propto f_1(x^h|x^s)f_0(x^s).$$

This resulting joint distribution $f_1(x^h|x^s)f_0(x^s)$ is clearly counterfactual because the conditional health-status is for the minority, but the marginal SES is for the majority.

Generalization

We can generalize the weights for the ATE and ATT as follows:

$$\omega_1(x) = \frac{\omega_1^B(x)}{\omega_1^B(x^s)}, \text{ and } \omega_0(x) = \frac{\omega_0^B(x)}{\omega_0^B(x^s)}. \tag{8}$$

For the ATE, $\{\omega_1^B(x), \omega_0^B(x)\} = \{1/e(x), 1/(1 - e(x))\}$ and $\{\omega_1^B(x^s), \omega_0^B(x^s)\} = \{1/e(x^s), 1/(1 - e(x^s))\}$. For the ATT, $\{\omega_1^B(x), \omega_0^B(x)\} = \{1, e(x)/(1 - e(x))\}$ and $\{\omega_1^B(x^s), \omega_0^B(x^s)\} = \{1, e(x^s)/(1 - e(x^s))\}$.

The weighted minority distribution by $\omega_1(x)$ is

$$f_1(x)\omega_1(x) = f_1(x) \frac{g(x)}{e(x)} \frac{e(x^s)}{g(x^s)} \propto f(x)e(x^s) \frac{g(x)}{g(x^s)} \propto \left\{ f(x^h|x^s) \frac{g(x)}{g(x^s)} \right\} f_1(x^s).$$

In the same way, we can show that the weighted majority distribution by $\omega_0(x)$ is

Previously, we let $f_g(x^h|x^s) \equiv f(x^h|x^s)g(x)/g(x^s)$ define the target conditional distribution of health status. Thus, the proposed weights realize the target populations, expressed as $f_g(x^h|x^s)f_1(x^s)$ and $f_g(x^h|x^s)f_0(x^s)$, for minority and majority groups, respectively.

An important example arises when $g(x) = g(x^s) = 1$ for the ATE, where the target distribution of conditional health-status is easily obtained as $f(x^h|x^s)$. For the ATT, $g(x) = e(x)$ and $g(x^s) = e(x^s)$, and therefore, the target conditional distribution of health-status is

which is of the minority group.

We can demonstrate that the marginal distributions of health status variables are not necessarily the same between the groups in the hypothetical populations

generated by the proposed weights:

$$\int \frac{f_g(x^h|x^s)f_1(x^s)}{E[f_g(x^h|x^s)f_1(x^s)]} dx^s \neq \int \frac{f_g(x^h|x^s)f_0(x^s)}{E[f_g(x^h|x^s)f_0(x^s)]} dx^s.$$

The above inequality holds unless health status and SES variables are independent or two racial groups have the same SES distribution. Therefore, we do not focus on balancing the marginal distributions of health status. Instead, we seek to balance the conditional distributions of health status given SES.

The ATO has been used as an estimand to address lack of overlap [15], which often occurs when some covariates are highly correlated with the treatment variable. To identify the ATO measure of racial disparity, we let $g(x) = e(x)\{1 - e(x)\}$ and $g(x^s) = e(x^s)\{1 - e(x^s)\}$. Then, the corresponding overlap weights are $\omega_1(x) = \{1 - e(x)\}/\{1 - e(x^s)\}$ and $\omega_0(x) = e(x)/e(x^s)$. Notably, the target population for the original overlap weights is

$$f(x)e(x)\{1 - e(x)\} \propto f_1(x)f_0(x)/f(x), \tag{9}$$

which is the product of the minority and majority densities divided by the marginal density of the combined. We can show that the target conditional distribution of health status has the same form as Eq. (9):

$$f_g(x^h | x^s) = f(x^h | x^s) \frac{e(x)\{1 - e(x)\}}{e(x^s)\{1 - e(x^s)\}} \propto \frac{f_1(x)f_0(x)/f(x)}{f_1(x^s)f_0(x^s)/f(x^s)} = \frac{f_1(x^h | x^s)f_0(x^h | x^s)}{f(x^h | x^s)}$$

Therefore, the target conditional distribution is the overlapped conditional distribution of health status given SES between the groups.

Estimation of health disparity measures

Based on Eqs. (4) and (5), we can construct consistent estimators for μ_{g1} and μ_{g0} . Using our result that $f_1(x)\omega_1(x) \propto f_g(x^h | x^s)f_1(x^s)$, we can write μ_{g1} in Eq. (4) as

$$\mu_{g1} = \frac{\int E[Y | Z = 1, X = x]f_1(x)\omega_1(x)dx}{\int f_1(x)\omega_1(x)dx}.$$

Using our result that $f_0(x)\omega_0(x) \propto f_g(x^h | x^s)f_0(x^s)$, we can also write μ_{g0} in Eq. (5) as

$$\mu_{g0} = \frac{\int E[Y | Z = 0, X = x]f_0(x)\omega_0(x)dx}{\int f_0(x)\omega_0(x)dx}.$$

Therefore, we use the following consistent estimator for τ_g :

$$\hat{\tau}_g = \frac{\sum_i \omega_1(X_i)Z_i Y_i}{\sum_i \omega_1(X_i)Z_i} - \frac{\sum_i \omega_0(X_i)(1-Z_i)Y_i}{\sum_i \omega_0(X_i)(1-Z_i)}. \tag{10}$$

Estimation of the propensity scores

In the proposed framework, the important task is to balance the conditional distributions of health status given SES between minority and majority groups. Usual methods for PS estimation may not be effective in achieving this task. However, we can view this task as estimating the PSs within each subgroup defined by SES variables,

where the PSs are intended to balance health-status variables. Several researchers have investigated PS methods for subgroup analysis. Dong et al. [16] demonstrated that fitting the PS models by strata yields better covariate balance within the strata than fitting the main-effects model on the whole data. Therefore, we propose to fit the full PS model to each subgroup defined by all SES variables. As shown by Yang et al. [17], the overlap weights achieve exact balance within each subgroup when the PS models are fitted by strata and maximum likelihood (ML) is used for parameter estimation. However, exact balance is not guaranteed for the other weights, such as the IPW weights, when ML is used for parameter estimation. In their simulation study, Dong et al. [16] showed that the covariate balancing propensity score (CBPS) [18] is effective in balancing the covariates within strata when it is

used to estimate the stratum-specific PS models. To have a better finite balance for the weights other than the overlap weights, we use CBPS for PS estimation.

Measuring concordance with IOM's definition

In typical PS-weighted analysis, we can use the absolute standardized mean difference (ASMD) [19] to evaluate whether the covariates are balanced between groups in the target population. If X_{ij} indicates covariate j of subject i , then the ASMD of X_{ij} is

$$D(X_{ij}) = \frac{\left| \frac{\sum_{i=1}^N X_{ij}Z_i\omega_1^B(X_i)}{\sum_{i=1}^N Z_i\omega_1^B(X_i)} - \frac{\sum_{i=1}^N X_{ij}(1-Z_i)\omega_0^B(X_i)}{\sum_{i=1}^N (1-Z_i)\omega_0^B(X_i)} \right|}{\sqrt{\frac{s_{1(j)}^2 + s_{0(j)}^2}{2}}},$$

where $s_{1(j)}^2$ is the sample variance of unweighted X_{ij} in the minority group, and $s_{0(j)}^2$ is the sample variance of unweighted X_{ij} in the majority group.

Following IOM's definition, we do not balance all covariates, pursued by typical causal inference approaches. Therefore, we need to modify $D(X_{ij})$ for health disparities research. First, we adopt the metrics introduced by Choi et al. [8] to measure the degree to which the proposed weights alter the minority SES:

$$D_1(X_{ij}^s) = \frac{\left| \frac{\sum_{i=1}^N X_{ij}^s Z_i}{\sum_{i=1}^N Z_i} - \frac{\sum_{i=1}^N X_{ij}^s Z_i \omega_1(X_i)}{\sum_{i=1}^N Z_i \omega_1(X_i)} \right|}{s_{1(j)}}, \tag{11}$$

where $j = 1, \dots, J$. The metric to measure the degree to which the proposed weights alter the majority SES is:

$$D_0(X_{ij}^s) = \frac{\left| \frac{\sum_{i=1}^N X_{ij}^s (1-Z_i) - \sum_{i=1}^N X_{ij}^s (1-Z_i) \omega_0(X_i)}{\sum_{i=1}^N (1-Z_i)} - \frac{\sum_{i=1}^N X_{ij}^s (1-Z_i) \omega_0(X_i)}{\sum_{i=1}^N (1-Z_i) \omega_0(X_i)} \right|}{s_{0(j)}}, \tag{12}$$

where $j = 1, \dots, J$.

We demonstrated that the marginal distributions of health status are generally not balanced if the weights preserve the original SES distributions. Thus, we seek to measure the extent to which health-status variables are balanced within the SES subgroups. Without loss of generality, assume that the SES variables are categorical, and the total number of subgroups that can be generated by them is $R = 2^K$. In our application to the RHC data, we consider four subgroups defined based on two binary SES variables, education and income. Let $G_i \in \{1, \dots, R\}$ denote the subgroup indicator for subject i . Define an indicator function $1(A)$, equal to 1 if statement A is true and 0 otherwise. Adopting Eq. (6) of Yang et al. [17], we modify $D(X_{ij})$ to measure the balance in the health-status variables in subgroup r :

$$D_r(X_{ij}^h) = \frac{\left| \frac{\sum_{i=1}^N X_{ij} 1(G_i=r) Z_i \omega_1(X_i)}{\sum_{i=1}^N 1(G_i=r) Z_i \omega_1(X_i)} - \frac{\sum_{i=1}^N X_{ij} 1(G_i=r) (1-Z_i) \omega_0(X_i)}{\sum_{i=1}^N 1(G_i=r) (1-Z_i) \omega_0(X_i)} \right|}{\sqrt{\frac{s_{r1(j)}^2 + s_{r0(j)}^2}{2}}}, \tag{13}$$

where $j = J + 1, \dots, J + K$, $r = 1, \dots, R$, $s_{r1(j)}^2$ is the sample variance of unweighted X_{ij} in the minority group belonging to SES subgroup r , and $s_{r0(j)}^2$ is the sample variance of unweighted X_{ij} in the majority group belonging to SES subgroup r .

In this section, we have discussed three ASMD measures, $D_1(X_{ij}^s)$, $D_0(X_{ij}^s)$, and $D_r(X_{ij}^h)$, to evaluate whether given weights create target populations concordant with IOM’s definition. Smaller values of these three indicate that the given weights are reliable. Austin and Stuart [19] recommended using 0.1 as a threshold to detect covariate imbalance. Following this recommendation, if there are any covariates whose ASMD values are greater than 0.1, this will indicate that the weights may not achieve the IOM’s definition successfully.

Characterizing target populations

It would be important to characterize the target populations identified by the proposed weights because they are usually hypothetical and not directly observable from the data. Adopting the estimators for μ_{g1} and μ_{g0} in (10), we can estimate the means of X_{ij} in the target joint distributions. The estimator for the mean of X_{ij} the weighted minority group is

$$m_1(X_{ij}) = \frac{\sum_i \omega_1(X_i) Z_i X_{ij}}{\sum_i \omega_1(X_i) Z_i}, \tag{14}$$

and that for the weighted majority group is

$$m_0(X_{ij}) = \frac{\sum_i \omega_0(X_i) (1-Z_i) X_{ij}}{\sum_i \omega_0(X_i) (1-Z_i)}. \tag{15}$$

The estimators $m_1(X_{ij})$ and $m_0(X_{ij})$ are obtained by replacing outcome Y_i with covariate X_{ij} in those of μ_{g1} and μ_{g0} .

Application to right heart catheterization data

Connors et al. [10] examined the effectiveness of RHC on critically ill patients during the first 24 h after admission to the intensive care unit. The authors compared survival outcomes and found that patients managed with RHC had poorer survival outcomes than those without RHC. The original study [10] and other studies [9, 20, 21] focused on whether RHC was beneficial or harmful for critically ill patients. In this study, we examined whether there was a disparity in receiving RHC between black and white patients admitted to the intensive care unit. If there is a difference in the probabilities of receiving RHC even after adjusting for the variables relating to clinical needs and health status, while keeping the original SES distributions, that difference would indicate some racial disparities in receiving RHC. Because physicians believed that management with RHC leads to better patient outcomes at the time of the RHC study [10], even though the actual RHC study demonstrated that it was not beneficial in general, reduced use of RHC by black patients would reflect disparities. We compared the unadjusted proportions of patients receiving RHC, which indicated that black patients were 1.86% less likely to receive RHC than white patients. This proportion difference was not statistically significant, with a p-value of 0.286 at a 5% level. However, this estimate did not account for the differences in clinical needs and health status, which should not contribute to health disparity measures.

The original data includes 920 blacks, 4460 whites, and 335 others, but we considered only black and white patients to illustrate our methods. Thus, our data consisted of 5380 patients. We considered twenty continuous (age, number of comorbid illnesses, and baseline laboratory values) and four categorical (sex, primary disease category, cancer, and ‘do not resuscitate’ (DNR) status) variables for health status. After generating the dummy variables for the categorical variables, we obtained twenty-seven health-status variables. Table S1 provides a full list of the health-status variables. For SES, we considered years of education and income. We converted

Table 2 Joint distribution of education and income in right heart catheterization data

Education	Income	
	Less than \$25K	\$25K or greater
Less than high school	3321	532
High school or more	801	726

years of education to a binary variable, equal to 1 if high school education or more and 0 otherwise. Income had four categories, but these categories were pooled to have the following two categories: 'less than \$25K' and '\$25K or greater'. Table 1 presents the baseline characteristics of non-RHC and RHC patients by race. Table 2 shows the frequency distribution of the SES subgroups.

In the context of this analysis, the PS was defined as the probability of being a black patient conditional on the covariates. The goal of this analysis was to demonstrate the use of the deweighting approach to estimate the ATE, ATT, and ATO measures of health disparities in receiving RHC. The full PS model for the deweighting approach included the main effects of all health-status variables but was estimated within each of the SES subgroups, using the CBPS. Thus, the full PSs were obtained from these SES Stratum-specific Health-Status (SSHS)-PS models. The SES-PS model was fitted with the main effects of education and income, and their interaction term. Because the SES variables were categorical, the model was saturated. Finally, the proposed weights for the ATE, ATT, and ATO were calculated using the estimated full PSs and SES-PSs.

Based on Eq. (10), the proposed weights were used to estimate the ATE, ATT, and ATO measures of racial

disparities. We used the survey R package [22] to calculate 95% confidence intervals and presented inference results in Table 3. We used four methods to estimate each disparity measure: (i) full PS using the main effects of health-status and SES variables, (ii) full PS based on SSHS-PS, (iii) health-status PS, and (iv) deweighting. These full or health-status PSs were estimated using the CBPS. Note that methods (i) and (ii) used the PSs adjusting for all covariates, and methods (iii) and (iv) adjusted for only the health-status variables. Even though all estimates were not statistically significant at a 5% level, we observed meaningful differences between the methods regarding point estimates. The ATT estimates based on the full PS and SSHS-PS were positive, indicating that black patients were more likely to receive RHC by 1.51% and 0.64%, respectively. Therefore, using all covariates over-adjusted the crude difference, and as a result, the sign of the estimate was reversed. The ATT estimates based on the health-status PS and deweighting reduced the disparities to -0.14% and -0.35%, respectively, from -1.86%. Similar observations were found for the ATO estimates. For the ATE, only the deweighting estimate (-2.44%) represented a larger disparity than the unadjusted estimate (-1.86%).

We used our three metrics in Eqs. (11, 12, 13) to examine how better the proposed weights achieved the IOM's definition than the balancing weights based on the SSHS-PS, which we will call SSHS weights. Note that the proposed weights are obtained as the SSHS weights divided by the balancing weights based on the SES-PS. First, we checked whether the marginal SES distributions of blacks and whites were preserved in the ATE weighted data. To this end, we looked at whether the values of $D_0(X_{ij}^s)$ were sufficiently small for education and income. The values of

Table 3 Racial disparity estimates (%), as measured by the average controlled differences, in receipt of a right heart catheterization

Estimand	Method	Point estimate	Lower bound	Upper bound
NA	Unadjusted difference	-1.86	-5.25	1.59
ATE	Full PS	-0.16	-4.70	4.38
	SES stratum-specific health-status PS	-1.33	-6.27	3.62
	Health-status PS	-0.03	-3.75	3.69
	Deweighting	-2.44	-6.65	1.77
ATT	Full PS	1.51	-2.28	5.30
	SES stratum-specific health-status PS	0.64	-3.13	4.41
	Health-status PS	-0.14	-3.76	3.47
	Deweighting	-0.35	-4.12	3.41
ATO	Full PS	0.98	-2.69	4.64
	SES stratum-specific health-status PS	0.27	-3.38	3.92
	Health-status PS	-0.48	-4.05	3.08
	Deweighting	-0.88	-4.54	2.77

NA Not available, ATE Average treatment effect, ATT Average treatment effect in the treated, ATO Average treatment effect in the overlap population, PS Propensity score, SES Socioeconomic status

$D_0(X_{ij}^s)$ for education and income were (0.020, 0.050) when the proposed weights were used and (0.024, 0.008) when the SSHS weights were used. The values of $D_1(X_{ij}^s)$ for education and income were (0.019, 0.035) when the proposed weights were used and (0.275, 0.426) when the SSHS weights were used. Therefore, the black distribution of SES was significantly changed by the SSHS weights. Next, we investigated whether the values of $D_0(X_{ij}^s)$ were sufficiently small for education and income in the ATT weighted data. The values of $D_0(X_{ij}^s)$ were approximately 0.0003 for education and income when the proposed weights were used and (0.257, 0.324) with the SSHS weights. Thus, the white distribution of SES was significantly changed by the SSHS weights. All values of $D_1(X_{ij}^s)$ were 0 because a unit weight was applied to all black patients. Finally, we checked both $D_1(X_{ij}^s)$ and

$D_0(X_{ij}^s)$ in the ATO population. The values of $D_0(X_{ij}^s)$ were (0.004, 0.0001) with the proposed weights and (0.229, 0.290) with the SSHS weights. Those of $D_1(X_{ij}^s)$ were (0.008, 0.002) with the proposed weights and (0.088, 0.122) with the SSHS weights. These numerical results demonstrated that the proposed weights sufficiently preserved the original SES distributions for both groups in the ATE, ATT, and ATO populations, while the SSHS weights failed to do so.

In addition to the degree to which SES variables were altered, we checked the covariate balance of the health-status variables after the proposed weights were applied. Particularly, we investigated how CBPS and ML performed differently in terms of balancing health status variables when they were used to estimate the SSHS weights. In the ATE population, within each SES subgroup, we checked whether the health-status variables of

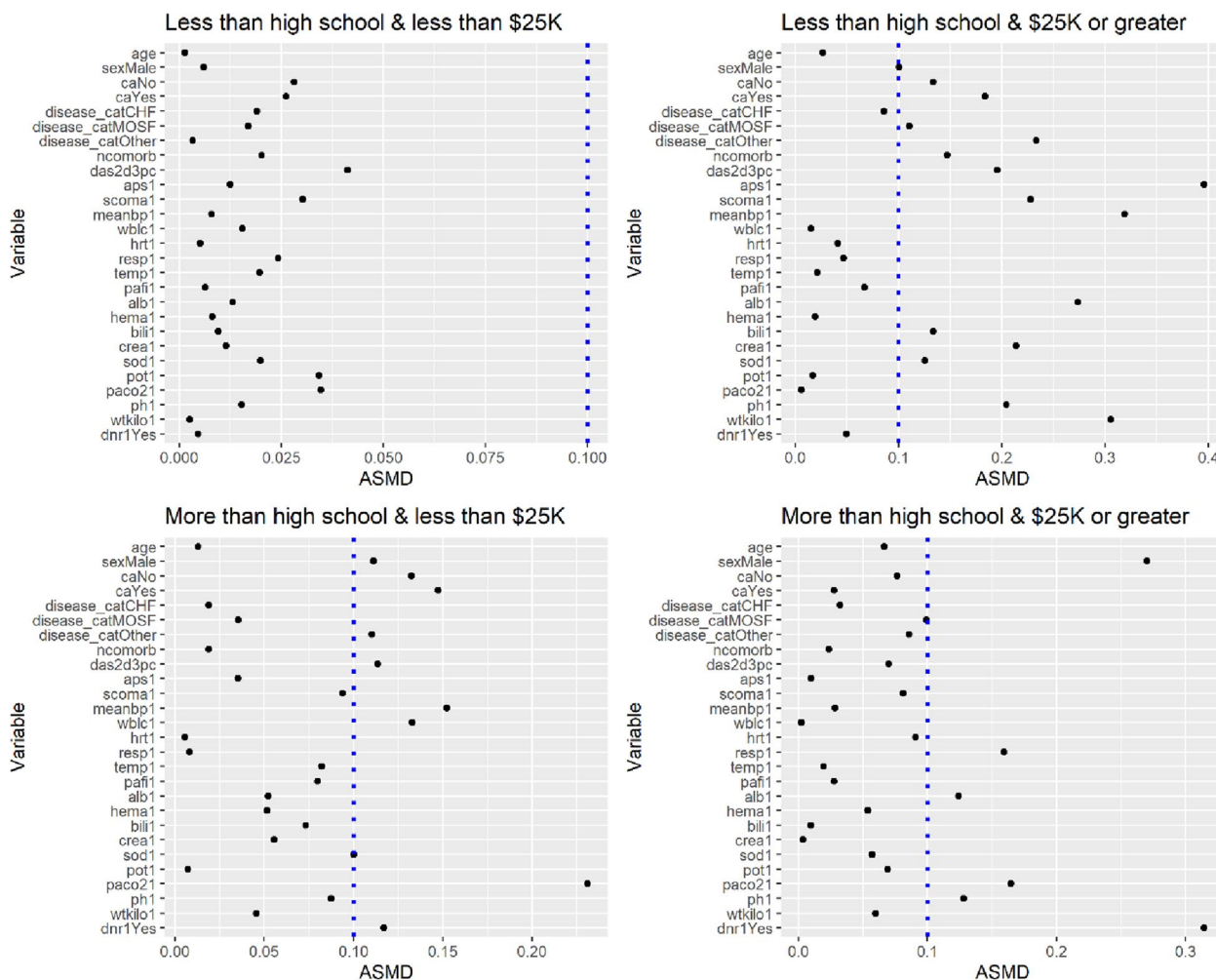


Fig. 1 Absolute standardized mean differences (Eq. 13) of the health-status variables in the combined group within each of the four subgroups defined by education and income in right heart catheterization data. The full propensity scores of black patients were estimated by maximum likelihood

blacks and whites were balanced. When ML was used, there was a severe imbalance in the health-status variables between groups: except for the SES subgroup that was educated less than high school and had an income of less than \$25K, many health-status variables had the values of $D_r(X_{ij}^h)$ over 0.1 (Fig. 1). However, when CBPS was used, all values of $D_r(X_{ij}^h)$ across the SES subgroups were smaller than 0.1. In the ATT population, within each of the SES subgroups, we checked whether the health-status variables of whites became similar to those of blacks after weighting. When ML was used, all values of $D_r(X_{ij}^h)$ for the health-status variables within each SES subgroup were smaller than 0.1, but a few values were near 0.1 (Fig. 2). For example, the values of $D_r(X_{ij}^h)$ for Glasgow Coma Score and temperature were 0.099 and 0.084, respectively, in the subgroup that was educated less than high school but had an income of \$25K or

greater. However, when CBPS was used, the covariate balance of all health-status variables in the ATT population was satisfactory with all values of $D_r(X_{ij}^h)$ almost 0. Those of all health-status variables in the ATO population were also almost 0.

We used the estimators $m_1(X_{ij})$ and $m_0(X_{ij})$ in Eqs. (14) and (15) to characterize black and white patients in the ATE, ATT, and ATO populations. Table 4 shows the characteristics of the ATE-weighted black and white patients, and the combined sample. The marginal means of the health-status variables between the weighted black and white patients were similar to those of the combined sample. The education and income of the weighted white patients were minimally altered. In the original samples, 18.6% of black patients had a high-school education or more, but 30.4% of white patients were so educated. After the proposed weighting, those percentages were changed to 19.3% and 31.3%, respectively. Regarding income,

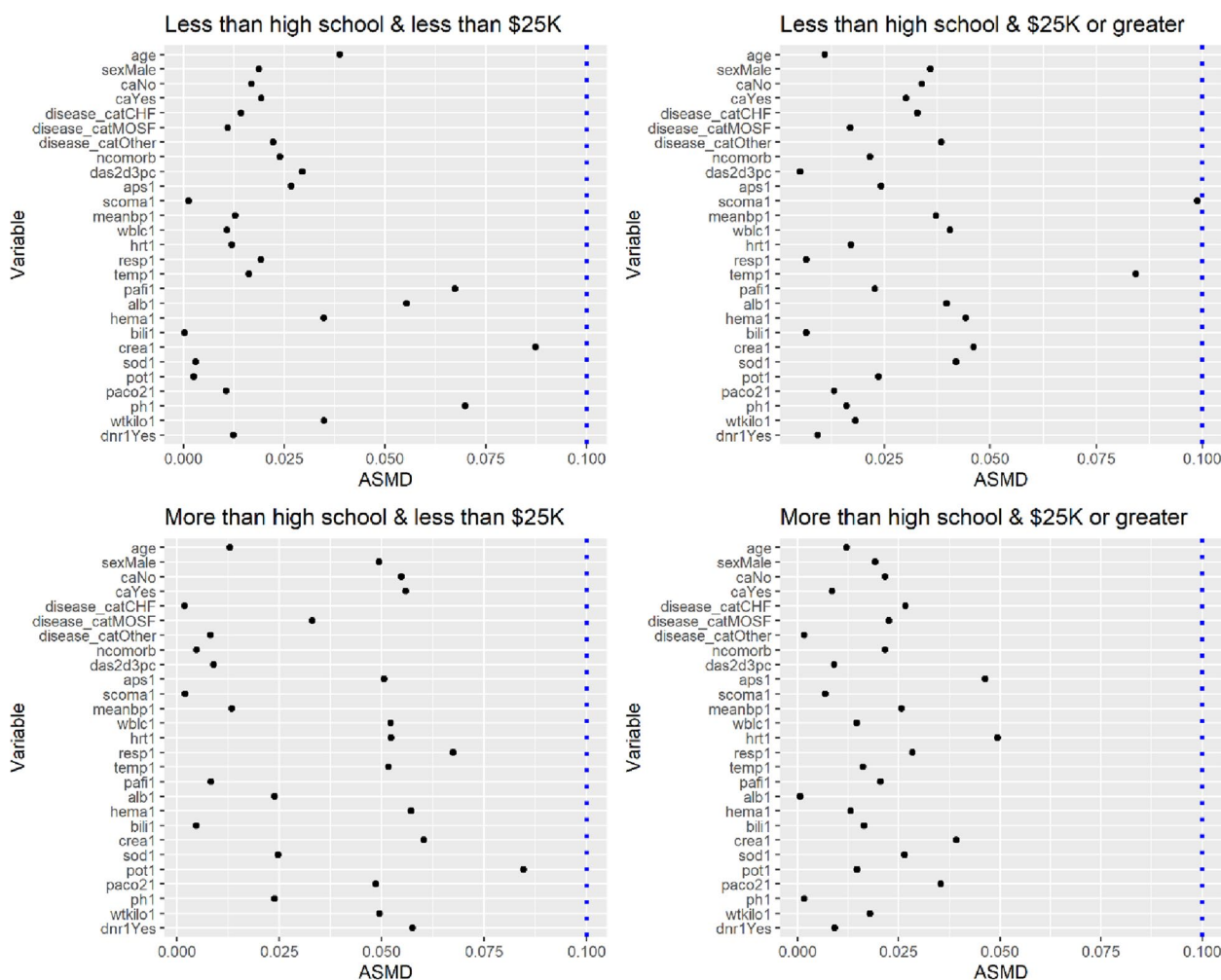


Fig. 2 Absolute standardized mean differences (Eq. 13) of the health-status variables in the black group within each of the four subgroups defined by education and income in right heart catheterization data. The full propensity scores of black patients were estimated by maximum likelihood

Table 4 Characteristics of the ATE-weighted black and white patients, and the original combined sample in right heart catheterization data

	Weighted black patients	Weighted white patients	Overall
N	940.74	4652.67	5380
Age (mean (SD))	62.28 (16.75)	61.04 (16.79)	61.99 (16.46)
Sex = Male (%)	511.1 (54.3)	2552.5 (54.9)	2994 (55.7)
Cancer (%)			
Metastatic	62.2 (6.6)	353.7 (7.6)	362 (6.7)
No cancer	721.8 (76.7)	3506.4 (75.4)	4105 (76.3)
Localized	156.7 (16.7)	792.6 (17.0)	913 (17.0)
Disease category (%)			
ARF	410.5 (43.6)	2029.8 (43.6)	2354 (43.8)
CHF	71.7 (7.6)	353.5 (7.6)	422 (7.8)
MOSF	260.2 (27.7)	1343.8 (28.9)	1504 (28.0)
Other	198.4 (21.1)	925.6 (19.9)	1100 (20.4)
Number of comorbidities (mean (SD))	1.53 (1.16)	1.49 (1.15)	1.51 (1.16)
Duke activity status index (mean (SD))	20.23 (5.36)	20.59 (5.36)	20.48 (5.31)
APACHE score (mean (SD))	54.98 (19.91)	55.10 (20.06)	54.59 (19.88)
Glasgow coma score (mean (SD))	21.78 (29.75)	21.76 (31.25)	21.03 (30.48)
Mean blood pressure (mean (SD))	79.26 (40.22)	78.91 (38.59)	78.56 (38.16)
WBC (mean (SD))	15.85 (13.10)	15.64 (11.65)	15.77 (11.98)
Heart rate (mean (SD))	114.17 (43.36)	114.92 (41.87)	114.98 (41.52)
Respiratory rate (mean (SD))	27.96 (14.94)	27.93 (14.06)	28.05 (14.12)
Temperature (mean (SD))	37.59 (1.79)	37.65 (1.78)	37.61 (1.78)
PaO ₂ /FiO ₂ ratio (mean (SD))	223.90 (111.31)	223.53 (120.81)	221.31 (114.15)
Albumin (mean (SD))	3.08 (0.65)	3.07 (0.81)	3.09 (0.79)
Hematocrit (mean (SD))	32.01 (8.43)	31.76 (8.38)	31.95 (8.38)
Bilirubin (mean (SD))	2.05 (4.18)	2.13 (4.41)	2.11 (4.37)
Creatinine (mean (SD))	2.20 (2.06)	2.19 (2.18)	2.13 (2.05)
Sodium (mean (SD))	136.91 (7.52)	136.83 (7.67)	136.81 (7.63)
Potassium (mean (SD))	4.08 (1.02)	4.07 (1.03)	4.07 (1.02)
PaCO ₂ (mean (SD))	39.11 (13.28)	38.62 (13.29)	38.80 (13.15)
PH (mean (SD))	7.38 (0.10)	7.39 (0.11)	7.39 (0.11)
Weight (mean (SD))	68.45 (29.27)	68.36 (29.22)	68.25 (29.05)
'Do not resuscitate' status on day 1 = Yes (%)	112.1 (11.9)	508.4 (10.9)	626 (11.6)
High school education or more (%)	182.0 (19.3)	1457.2 (31.3)	1527 (28.4)
Income >=\$25k (%)	120.1 (12.8)	1301.7 (28.0)	1258 (23.4)

11.6% of black patients earned \$25K or greater, while 25.8% of white patients earned so. After the proposed weighting, those percentages were changed to 12.8% and 28.0%, respectively. Table 5 shows the original covariates of black patients, and the ATT-weighted and original covariates of white patients. The marginal means of the health-status variables between the black and weighted white patients were similar. The education and income of the weighted white patients were completely unaltered. As a result, the original differences in these SES variables were preserved. Table 6 shows the ATO-weighted covariates in black and white patients. Again, we observed that

education and income were not altered, and the marginal distributions of the health-status variables were similar between the weighted groups.

Discussion

In our empirical example, to balance health status conditional on SES, we fitted the full PS model to each of the four strata based on education and income. However, if the number of SES subgroups is large, then fitting the full PS models by the SES strata can produce unstable PS estimates, which results in poor balance of health status. This problem can occur if some strata sizes are too

Table 5 Characteristics of black, the ATT-weighted, and original white patients in right heart catheterization data

	Black patients	Weighted white patients	White patients
N	920	4461.97	4460
Age (mean (SD))	56.50 (17.50)	55.01 (17.78)	63.13 (16.00)
Sex= Male (%)	455.0 (49.5)	2215.0 (49.6)	2539 (56.9)
Cancer (%)			
Metastatic	67.0 (7.3)	357.3 (8.0)	295 (6.6)
No cancer	742.0 (80.7)	3563.5 (79.9)	3363 (75.4)
Localized	111.0 (12.1)	541.2 (12.1)	802 (18.0)
Disease category (%)			
ARF	376.0 (40.9)	1865.8 (41.8)	1978 (44.3)
CHF	79.0 (8.6)	389.9 (8.7)	343 (7.7)
MOSF	286.0 (31.1)	1412.9 (31.7)	1218 (27.3)
Other	179.0 (19.5)	793.4 (17.8)	921 (20.7)
Number of comorbidities (mean (SD))	1.39 (1.13)	1.34 (1.11)	1.54 (1.16)
Duke activity status index (mean (SD))	20.58 (5.61)	20.97 (5.58)	20.46 (5.25)
APACHE score (mean (SD))	56.30 (20.92)	56.12 (20.60)	54.24 (19.64)
Glasgow coma score (mean (SD))	24.45 (31.87)	23.72 (32.46)	20.33 (30.14)
Mean blood pressure (mean (SD))	85.34 (42.59)	85.71 (40.84)	77.16 (37.04)
WBC (mean (SD))	16.12 (13.66)	16.21 (12.34)	15.70 (11.60)
Heart rate (mean (SD))	115.61 (43.64)	117.06 (41.78)	114.85 (41.08)
Respiratory rate (mean (SD))	28.68 (15.28)	28.83 (14.16)	27.93 (13.87)
Temperature (mean (SD))	37.46 (1.95)	37.54 (1.92)	37.65 (1.74)
PaO ₂ /FiO ₂ ratio (mean (SD))	244.08 (122.02)	241.38 (130.85)	216.62 (111.90)
Albumin (mean (SD))	3.04 (0.69)	3.04 (0.92)	3.10 (0.81)
Hematocrit (mean (SD))	30.50 (8.39)	30.52 (8.31)	32.25 (8.34)
Bilirubin (mean (SD))	2.12 (4.33)	2.12 (4.10)	2.11 (4.37)
Creatinine (mean (SD))	2.63 (2.71)	2.54 (2.76)	2.03 (1.87)
Sodium (mean (SD))	137.56 (7.74)	137.35 (8.00)	136.66 (7.59)
Potassium (mean (SD))	4.12 (1.12)	4.14 (1.11)	4.06 (1.00)
PaCO ₂ (mean (SD))	37.74 (12.46)	37.33 (12.71)	39.02 (13.28)
PH (mean (SD))	7.38 (0.11)	7.39 (0.11)	7.39 (0.11)
Weight (mean (SD))	68.92 (29.87)	69.18 (29.98)	68.11 (28.88)
'Do not resuscitate' status on day 1 =Yes (%)	79.0 (8.6)	347.9 (7.8)	547 (12.3)
High school education or more (%)	171.0 (18.6)	1356.0 (30.4)	1356 (30.4)
Income >=\$25k (%)	107.0 (11.6)	1151.0 (25.8)	1151 (25.8)

small, or extreme PSs exist. In such a case, more sophisticated methods than the subgroup fits would be needed. For example, Dong et al. [16] devised a stochastic search algorithm that selects one of two competing models for each subgroup, where model comparisons are performed based on covariate balance. Yang et al. [17] demonstrated that LASSO [23] is effective in reducing the number of interactions between the covariates to be balanced and strata variables for subgroup analysis of ATO. Validating the use of these methods for the deweighting approach merits further investigation.

As in traditional associational or causal analyses, the interpretation and validity of our approach depend on

what data are measured. Both the health-status PS and deweighting approaches commonly adjust for health-status variables, and thus, the resultant estimates represent racial differences when the observed health status or clinical needs are balanced toward a specific target population. However, these approaches are different in terms of how unmeasured SES variables affect their interpretations and validities. In the health-status PS approach, the PS model is fitted using only the health-status variables, and SES variables are not directly involved in the estimation of the PSs. This approach theoretically captures all measured and unmeasured SES variables in the analysis. However, it is impossible to assess how the unobserved

Table 6 Characteristics of the ATO-weighted black and white patients in right heart catheterization data

	Weighted black patients	Weighted white patients
N	823.21	3979.7
Age (mean (SD))	58.38 (17.16)	56.80 (17.36)
Sex = Male (%)	417.4 (50.7)	2019.3 (50.7)
Cancer (%)		
Metastatic	59.1 (7.2)	312.3 (7.8)
No cancer	652.3 (79.2)	3127.5 (78.6)
Localized	111.8 (13.6)	539.9 (13.6)
Disease category (%)		
ARF	346.0 (42.0)	1700.0 (42.7)
CHF	70.7 (8.6)	346.6 (8.7)
MOSF	242.9 (29.5)	1205.9 (30.3)
Other	163.6 (19.9)	727.2 (18.3)
Number of comorbidities (mean (SD))	1.43 (1.14)	1.38 (1.12)
Duke activity status index (mean (SD))	20.44 (5.57)	20.83 (5.47)
APACHE score (mean (SD))	55.47 (20.55)	55.43 (20.53)
Glasgow coma score (mean (SD))	23.40 (30.99)	22.99 (31.91)
Mean blood pressure (mean (SD))	83.39 (41.54)	83.68 (40.02)
WBC (mean (SD))	15.94 (13.30)	15.88 (11.89)
Heart rate (mean (SD))	115.38 (43.53)	116.74 (41.58)
Respiratory rate (mean (SD))	28.47 (15.22)	28.64 (14.14)
Temperature (mean (SD))	37.52 (1.89)	37.59 (1.85)
PaO ₂ /FiO ₂ ratio (mean (SD))	235.88 (117.25)	234.07 (122.65)
Albumin (mean (SD))	3.05 (0.68)	3.05 (0.88)
Hematocrit (mean (SD))	31.08 (8.40)	31.01 (8.25)
Bilirubin (mean (SD))	2.09 (4.30)	2.12 (4.15)
Creatinine (mean (SD))	2.37 (2.30)	2.33 (2.40)
Sodium (mean (SD))	137.31 (7.57)	137.13 (7.85)
Potassium (mean (SD))	4.08 (1.06)	4.09 (1.07)
PaCO ₂ (mean (SD))	38.35 (12.78)	37.86 (13.02)
PH (mean (SD))	7.38 (0.11)	7.39 (0.11)
Weight (mean (SD))	68.70 (29.94)	68.89 (29.82)
'Do not resuscitate' status on day 1 = Yes (%)	74.9 (9.1)	322.3 (8.1)
High school education or more (%)	150.3 (18.3)	1201.9 (30.2)
Income >=\$25k (%)	95.7 (11.6)	1029.1 (25.9)

SES variables are affected by weighting. In the dewatering approach, the SES variables are directly involved in preserving the SES distributions in comparison groups. Therefore, it can efficiently protect the observed SES distributions from being altered by weighting. However, this theoretically focuses only on the measured SES variables, which excludes the existence of potential unmeasured SES variables.

It must be beneficial to investigate other relevant definitions of health disparities. For example, the World Health Organization (WHO) defines health inequities as

“unfair, avoidable, or systematic differences in the health status and access to health resources of different population groups, whether those groups are defined socially, economically, demographically, or geographically or by other dimensions of inequality (e.g., sex, gender, ethnicity, disability, or sexual orientation)” [24]. The WHO also states that “health and health inequity are determined by the conditions in which people are born, grow, live, work, play and age, as well as biological determinants” [24]. Therefore, the WHO definition includes a broader range of SES factors as determinants of health inequities.

The proposed approach could implement the WHO definition by using more extensive SES measurements, including sex, elderly, housing, occupation, and disability, to generate the SES subgroups and SES-PS. Numerical evaluations of the deweighting approach using racial disparity data with comprehensive SES measurements merit future investigation.

Conclusion

In this study, we proposed a novel PS approach for estimating health disparities concordant with IOM's definition. The existing PS methods suffer from the critical limitation that SES variables cannot be preserved reliably. The approach overcomes this limitation by deweighting the subjects by SES-based PSs. We formally defined the estimands of health disparities and presented consistent estimators based on the proposed weights. Our analysis of RHC data demonstrated that the proposed deweighting is an effective method to estimate the measures of health care disparities.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12874-024-02230-5>.

Supplementary Material 1.

Acknowledgements

The author is grateful to the editor and three anonymous reviewers who provided valuable suggestions for improving the original submission of this paper.

Authors' contributions

B.C. developed the methods, conducted the study, and wrote the manuscript.

Funding

This work was partially supported by the National Cancer Institute for the Mays Cancer Center (P30CA054174) at the UT Health San Antonio. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The dataset used during the current study is available from <https://hbiostat.org/data/>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

- Jackson JW, VanderWeele TJ. Decomposition analysis to identify intervention targets for reducing disparities. *Epidemiology*. 2018;29:825–35.
- Ben-Michael E, Feller A, Kelz R, Keele L. Estimating Racial Disparities in Emergency General Surgery. 2023. arXiv preprint arXiv:2209.04321.
- Unequal Treatment. Confronting racial and ethnic disparities in health care (with CD). Washington, D.C: National Academies Press; 2003.
- Duan N, Meng X, Lin JY, Chen C, Alegria M. Disparities in defining disparities: statistical conceptual frameworks. *Stat Med*. 2008;27:3941–56.
- Cook BL, McGuire TG, Meara E, Zaslavsky AM. Adjusting for health status in non-linear models of health care disparities. *Health Serv Outcomes Res Methodol*. 2009;9:1–21.
- McGuire TG, Alegria M, Cook BL, Wells KB, Zaslavsky AM. Implementing the Institute of Medicine definition of disparities: an application to mental health care. *Health Serv Res*. 2006;41:1979–2005.
- Li F, Li F. Propensity score weighting for causal inference with multiple treatments. *Ann Appl Stat*. 2019;13:2389.
- Choi BY, Gelfond J, Kaushik D, Svatek RS, Wang C-P. Health status balancing weights for estimation of health care disparities. *Health Serv Outcomes Res Methodol*. 2023;23:185–206.
- Li F, Morgan KL, Zaslavsky AM. Balancing covariates via propensity score weighting. *J Am Stat Assoc*. 2018;113:390–400.
- Connors AF. The effectiveness of right heart catheterization in the initial care of critically ill patients. SUPPORT Investigators *JAMA*. 1996;276:889–97.
- Li F, Thomas LE, Li F. Addressing extreme propensity scores via the overlap weights. *Am J Epidemiol*. 2018. <https://doi.org/10.1093/aje/kwy201>.
- Li F, Zaslavsky AM, Landrum MB. Propensity score weighting with multi-level data. *Stat Med*. 2013;32:3373–87.
- Hirano K, Imbens GW, Ridder G. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*. 2003;71:1161–89.
- Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41–55.
- Crump RK, Hotz VJ, Imbens GW, Mitnik OA. Dealing with limited overlap in estimation of average treatment effects. *Biometrika*. 2009;96:187–99.
- Dong J, Zhang JL, Zeng S, Li F. Subgroup balancing propensity score. *Stat Methods Med Res*. 2020;29:659–76.
- Yang S, Lorenzi E, Papadogeorgou G, Wojdyla DM, Li F, Thomas LE. Propensity score weighting for causal subgroup analysis. *Stat Med*. 2021;40:4294–309.
- Imai K, Ratkovic M. Covariate balancing propensity score. *J R Stat Soc B*. 2014;76:243–63.
- Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statist Med*. 2015;34:3661–79.
- Hirano K, Imbens GW. Estimation of Causal effects using propensity score weighting: an application to data on right heart catheterization. *Health Serv Outcomes Res Methodol*. 2001;2:259–78.
- Choi BY. Subclassification estimation of the weighted average treatment effect. *Biom J*. 2021;63:1706–28.
- Lumley T. Analysis of complex survey samples. *J Stat Soft*. 2004;9:9.
- Tibshirani R. Regression shrinkage and selection via the lasso. *J Roy Stat Soc Ser B (Methodol)*. 1996;58:267–88.
- World Health Organization. Health equity. 2024. <https://www.who.int/health-topics/health-equity>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.