



Published in final edited form as:

*J Health Soc Behav.* 2022 June ; 63(2): 232–249. doi:10.1177/00221465211066108.

## Structural Racism and Quantitative Causal Inference: A Life Course Mediation Framework for Decomposing Racial Health Disparities

Nick Graetz<sup>1,2</sup>, Courtney E. Boen<sup>2,3</sup>, Michael H. Esposito<sup>4</sup>

<sup>1</sup>Department of Sociology, Princeton University

<sup>2</sup>Population Studies Center, University of Pennsylvania

<sup>3</sup>Department of Sociology, Population Aging Research Center, Leonard Davis Institute for Health Economics, University of Pennsylvania

<sup>4</sup>Department of Sociology, Washington University in St. Louis

### Abstract

Quantitative studies of racial health disparities often use static measures of self-reported race and conventional regression estimators, which critics argue is inconsistent with social constructivist theories of race, racialization, and racism. We demonstrate an alternative counterfactual approach to explain how multiple racialized systems dynamically shape health over time, examining racial inequities in cardio-metabolic risk in the National Longitudinal Study of Adolescent to Adult Health. This framework accounts for the dynamics of time-varying confounding and mediation that is required in operationalizing a “race” variable as part of a social process (*racism*) rather than a separable, individual characteristic. We decompose the observed disparity into three types of effects: a controlled direct effect (“unobserved racism”), proportions attributable to interaction (“racial discrimination”), and pure indirect effects (“emergent discrimination”). We discuss the limitations of counterfactual approaches while highlighting how they can be combined with critical theories to quantify how interlocking systems produce racial health inequities.

### Keywords

racism; racial health disparities; life course; mediation; g-computation

---

As early as Du Bois (1899), critical race scholars have linked racialized health disparities to the unequal social and economic conditions produced by racism. Over the past century, social scientific research has generally adopted and expanded a social constructionist view of race and racism. A large body of work describes the connections between systems of racial oppression, the construction of racial categories, and how exposure to racist systems becomes “embodied” to produce racialized health inequities (Bailey et al. 2017; Du Bois

---

**Corresponding Author** Nick Graetz, Department of Sociology, Princeton University, 106 Wallace Hall Princeton University, Princeton, NJ 08544. ngraetz@princeton.edu.

SUPPLEMENTAL MATERIAL

Additional supporting information may be found in the online version of this article.

1899; Geronimus et al. 2006; Goosby, Cheadle, and Mitchell 2018; Green and Darity 2010; Krieger 2018; Phelan and Link 2015; Reskin 2012; Sewell 2016).

Still, despite a strong theoretical foundation for a relational, social constructionist view of race and racism, quantitative analyses of racial health disparities generally fall short of translating theories of racism to empirical models beyond descriptive analyses (Muntaner 2013; Reskin 2012). Scholars have critiqued how “race” is interpreted in counterfactual frameworks, especially regression models (Kohler-Hausmann 2019; Reskin 2012; Sen and Wasow 2016; Williams 2019; Zuberi and Bonilla-Silva 2008). Particularly troubling is that, when applied to longitudinal data, conventional regression assumes an absence of time-varying relationships among variables. This assumption risks reifying race as a static individual trait that can be separated from other systems of social stratification such as social class position, rather than as mutually co-constituted with these other systems of inequality over time (Kohler-Hausmann 2019; Zuberi and Bonilla-Silva 2008). Critiques also stem from the limited set of counterfactuals offered by conventional regression models, where a race variable is often included in a model and additional variables are added to “explain away” racial variation in the outcome. Given the interpretation of the race variable as attempting to capture a relational process, however, this approach warrants critique. Is a static comparison of racial groups in which all other correlated exposures are held equal the most useful counterfactual we can estimate in quantitative studies? We argue that it is not.

What, then, are appropriate counterfactuals to quantify how a system of racism works to shape population health? And importantly, are these counterfactuals estimated by traditional regression models? We argue that conventional analyses to understand mechanisms that produce racialized variation in health require several strong, often untenable assumptions. By contrast, recent advancements in causal mediation analysis have provided new methods for modeling how dynamic social processes produce population health patterns over time (Bauer and Scheim 2019; Esposito 2019; Wang and Arah 2015). These approaches address several shortcomings of traditional quantitative approaches to studying racial disparities that can be of great utility to sociological research on racial health inequities by better accommodating dynamic and relational theories of racism and race.

In this study, we combine a novel technique for modelling the relative importance of multiple dependent mediators with a robust theory of the relational system of structural racism to examine life course processes contributing to racial health disparities. Given that the racial stratification of socioeconomic resources, risks, and opportunities is a key pathway linking racism to population health (Boen 2016; Phelan and Link 2015), we pay particular attention to the mediating, cumulative roles of socioeconomic exposures in the production of racial health disparities. Socioeconomic processes are not the only factors contributing to racialized health disparities but are an important and frequently studied component in mediation analyses. That socioeconomic exposures evolve across the life course is an important feature for demonstrating the utility of our approach for handling the complex interplay of time-varying mediators and confounders. Using longitudinal data from the National Longitudinal Study of Adolescent to Adult Health (Add Health), we decompose the total Black-white disparity in cardio-metabolic risk to three types of cumulative life course effects that each have intuitive interpretations (Jackson and VanderWeele 2019;

Wang and Arah 2015): 1) the controlled direct effect (CDE) of racism (i.e. unobserved pathways through which racism becomes embodied that do not operate through the observed mediators, which we call *unobserved racism*); 2) the portion attributable (PAI) to interaction for each observed socioeconomic mediator (i.e. how the effects of a socioeconomic exposure on health varies by racialized category, which we label *racial discrimination*); and 3) the pure indirect effect (PIE) for each observed mediator that we call *emergent discrimination* (Reskin 2012).

By combining a critical race theoretical approach with modern causal inference methods, our study highlights the limitations of conventional regression models for identifying the life course pathways of structural racism governing the production of racial health inequities. Importantly, we offer an alternative mediation framework for examining the social roots of racial health inequities that is more consistent with a social constructivist understanding of racism and race than conventional regression.

## Scope of Study

There are several productive points of entry for pushing quantitative research on health disparities closer to historical, contextual, relational theories of race, racism, and racialization (Bailey et al. 2017; Bailey, Feldman, and Bassett 2020; Diez Roux 2011; Ford and Airhihenbuwa 2010; Frerichs et al. 2016; Itzigsohn and Brown 2020; Krieger 2018; Laster Pirtle 2020; Sewell 2016; Williams 2019; Williams, Priest, and Anderson 2016). Here, we constrain our critique to problematize a specific inferential statement that arises in most contemporary regression-based analyses of racialized health disparities:

### **“X% of the racial disparity in outcome Y is explained by M (‘non-race’ variable).”**

We identify two specific problems with this notion: 1) issues of post-treatment bias — implied by a social-constructivist understanding of “race” as proxying time-varying systems of racism and racialization (Esposito 2019; Jackson and VanderWeele 2018; Naimi et al. 2016; Sen and Wasow 2016) — that plague the identification of “X%”; and 2) the conceptual separability of a “race” variable from all other variables that are nominally “not race” that is implied by common estimands (Kohler-Hausmann 2019; Sen and Wasow 2016; Zuberi and Bonilla-Silva 2008).

Our goal is not to make a new theoretical contribution about the systems of racism underlying health inequities; this exists elsewhere (e.g., Bailey et al. 2017; Bailey, Feldman, and Bassett 2020; Laster Pirtle 2020; Reskin 2012; Sewell 2016; Goosby et al. 2018; Krieger 2018). We also do not contend that there exists any singular quantitative design necessary to establish the causal role of racism in generating racialized health; this should be a starting point (Chowkwanyun and Reed 2020; Ford and Airhihenbuwa 2010; Williams and Sternthal 2010). Instead, we contend that as quantitative researchers continue to make use of regression models as the primary workhorse for examining racial health disparities from survey data, they must be clearer about the (often implicit) construction of causal graphs, target counterfactual contrasts, and general assumptions that underline the identification of

effects under study (Esposito 2019; Holland 2008; Lundberg, Johnson, and Stewart 2021; Sen and Wasow 2016). We explore one design for better translating social-constructivist theories of race and racism to empirical estimands, which may be particularly useful when using common individual-level, longitudinal survey data. We conclude with a discussion of which specific problems of conventional regression models we address by using modern counterfactual methods, while also highlighting how conceptual features of racism as a fundamental cause of health are still reduced within such methods (Phelan and Link 2015; Zuberi and Bonilla-Silva 2008).

## BACKGROUND

### The Social Construction of Race

In his essay “The Superior Race,” W.E.B. Du Bois (1940) imagines himself in a dialogue with a fictional white character, Roger Van Dieman. Du Bois tries to explain that race is a social, cultural, and historical fact – not a natural or inherent biological phenomenon. Van Dieman presses Du Bois, asking him how he can tell who is Black if race cannot be objectively measured or ascertained biologically. Du Bois closes the essay stating,

**“I recognize it quite easily and with full legal sanction: The Black man is a person who must ride the ‘Jim Crow’ in Georgia.”**

Du Bois’s retort carries tremendous meaning by concisely pointing to the dynamic and relational processes that serve to create, reify, and give meaning to race. A man is not Black because of some measurable individual trait or phenotypic characteristic; a man is Black because social, institutional, and legal structures *treat* him as Black. Since Du Bois, sociological scholarship has played a prominent role in expanding this social constructivist view of race, highlighting the essential roles of structures and institutions in shifting and maintaining racial boundaries (Bonilla-Silva 1997; Roberts 2011; Sewell 2016; Zuberi 2001; Zuberi, Patterson, and Stewart 2015).

Theoretically, recent studies of racial health inequities generally adopt a social-constructionist understanding of racism, racialization, and race, with a large body of work focused on the historical, social, and political roots of these disparities (Bailey et al. 2017; Phelan and Link 2015; Williams 2012). Still, identifying the role of structural racism in governing the distribution of various types of risk poses significant empirical challenges. A common approach in quantitative studies of racial health disparities is to include a time-invariant measure of race in regression models and subsequently control for correlated variables, such as socioeconomic status, to isolate variation that is explained by racial categories. This approach treats constructs like “race” and “socioeconomic position” as separable, rather than co-constituted and historically contingent, constructs. Still, structural racism has been described as “the totality of ways in which societies foster racial discrimination through mutually reinforcing systems of housing, education, employment, earnings, benefits, credit, media, health care, and criminal justice” (Bailey et al. 2017). The logic of conceptual separability implicit in regression estimators therefore carries a significant risk of reifying “race” as an independent construct, rather than part of the time-varying reciprocal process of racialization and racism within and across socioeconomic,

political, and cultural systems (Sewell 2016; Zuberi and Bonilla-Silva 2008; Zuberi et al. 2015).

### Studies Focusing on “Direct Effects” of Race

The approach to separating “race” from social exposures that are “not race” is the basic counterfactual logic used in audit and legal studies of racial discrimination. In cases involving hiring discrimination, for example, the residual racial variation that remains after considering other factors like educational attainment is interpreted as *the* effect of racial discrimination (Kohler-Hausmann 2019). This same logic is deployed in audit studies, which are appealing for the precision of their experimental manipulations. These studies are often interpreted as premier causal evidence of racial discrimination, as if assigning a stereotypically Black or white name to a hypothetical jobseeker, for example, while holding everything else constant, is capturing something that is more *truly representative of racism* than the complex, longitudinal systems that produce racialized distributions of everything else that might appear on a resume (and indeed, who is able to apply for which jobs).

While such studies identify important dynamics of explicit forms of racism, they typically offer no concurrent comparison of how much the total observed racial disparity in an outcome like hiring is due to employer decision-making at the final step before employment. We might also consider, for example, racial differences in exposure to penal systems that shape the probability of an employer receiving differently racialized individuals’ resumes in the first place. Indeed, by holding all else constant, audit studies implicitly suggest that it is possible to separate “racial discrimination” from factors that are “not race,” like socioeconomic position. But racial variation in *any* characteristic can only be the result of historical and contemporary projects of racism and racialization (Bonilla-Silva 2009; Roberts 2011; Williams et al. 2016).

### Structural Racism and the Race Discrimination System

Variation arising from racial stratification in the mutually reinforcing systems of labor, housing, and education is central to theories of structural racism (Bonilla-Silva 1997, 2009). Reskin’s (2012) description of the “race discrimination system” is useful for understanding the dynamic and relational processes producing racial health disparities. This framework highlights the need for a systems approach to describing the process of racial stratification as a broad causal system consisting of various subsystems (e.g., residential and school segregation, housing and mortgage markets, health services). Reskin (2012) contends that in the quantitative literature, scholars often attempt to isolate and manipulate a single subsystem and estimate the direct effect, assuming that doing so would not change other subsystems. In contrast, a systems approach to studying structural racism requires scholars to simultaneously consider that: 1) racialized disparities exist across many exposures, 2) disparities across exposures are mutually reinforcing, 3) the source of disparities is racial discrimination, and 4) there exists discrimination in the effects of exposures. Reskin’s (2012) use of the term “discrimination” can be confusing, given its inconsistent use across academic fields and colloquial interpretation. To the extent that there is *any* racial difference in an exposure-outcome association (i.e., an interactive effect that shows differential effects by race), the sub-system governing that exposure-outcome relationship is racializing

individuals and “discriminating” based on those racial categories. We will use this language of “*disparities* in exposures” and “*discrimination* in exposure effects” throughout this paper.

This longitudinal framework, which considers the causal dependence of multiple sub-systems of structural racism acting on racialized individuals over time, presents an array of methodological issues for quantitative inference, including concerns about mediation and confounding. Overcoming these issues is the primary purpose of this paper, as we believe they are important to addressing foundational concerns raised by Zuberi & Bonilla-Silva (2008), Kohler-Hausmann (2019), and others with regards to the implicit treatment of variables as independently “race” or “not race” in conventional regression.

### Traditional Approaches for Explaining Racial Disparities in Health

Link & Phelan (2015) describe racism as a fundamental cause of health disparities, a broad theory that is central to study of structural racism. The fundamental cause model is frequently invoked in quantitative studies of racial health disparities, typically paired with a regression model that includes mediating pathways. The fundamental cause model emphasizes the importance of meta-mechanisms, highlighting that the specific mechanisms through which the fundamental cause is operating at a given time can be easily substituted. Available data often inhibits directly studying the meta-mechanisms emphasized by the fundamental cause model, which is itself a considerable problem (Krieger 2018). Still, existing longitudinal cohort data can be usefully combined with cumulative life course theory and frameworks examining mediating pathways (Boen 2016; Brown et al. 2016; Kramer et al. 2017; Williams et al. 2016) to retrospectively study how a fundamental cause like racism operates through a particular space and time to produce racial health disparities.

However, the empirical challenges of considering and accounting for confounding—which have been described by Diez-Roux (2012), Reskin (2012), and Robinson and Bailey (2020)—are important to address in the context of mediation analyses. The most basic mediation model implied by Link & Phelan (2015) is illustrated with a directed-acyclic diagram (DAG) in Figure 1, which attempts to explain how much of the racial disparity in health ( $Y$ ) is explained by some indicator of material deprivation or socioeconomic position ( $M$ ). For illustrative purposes, consider that  $M$  is household income. This is a causal mediation question that relies on four key assumptions. As we will discuss below, each assumption is predicated on a specific understanding of how the exposure, typically operationalized as self-identified race, fits into the theoretical causal process of racism and racialization.

- **Assumption 1.** No unobserved confounding of  $\text{Race} \rightarrow Y$  ( $C$  contains all relevant confounding variables of race and health).
- **Assumption 2.** No unobserved confounding of  $\text{Race} \rightarrow M$  ( $C$  contains all relevant confounding variables of race and household income).
- **Assumption 3.** No unobserved confounding of  $M \rightarrow Y$  ( $L$  contains all relevant confounding variables of household income and health).



- **Assumption 4.** No  $M \rightarrow Y$  confounders ( $L$ ) affected by Race (the dotted arrow between Race and  $L$  does not exist; the “cross-world independence assumption”).

This generalized DAG is frequently combined with the following mediation formula for how much of the racial health disparity is “explained” by household income ( $M$ ), potentially including other “confounders” ( $L$ ,  $C$ ) in both models:

$$Y = \beta_0 + \beta_1(\text{Race}) + \beta_2(L) + \beta_3(C) + \varepsilon$$

$$Y = \alpha_0 + \alpha_1(\text{Race}) + \alpha_2(L) + \alpha_3(C) + \alpha_4(M) + \delta$$

**Proportion of racial disparity in  $Y$  explained by  $M = \frac{(\beta_1 - \alpha_1)}{\beta_1}$**

Often referred to as Baron-Kenny (1986) mediation, this approach simply asks: conditional on confounders, how much of the racial disparity in health outcome  $Y$  remains after controlling for the mediator  $M$  (income). There are three important considerations to discuss in this framework that are often neglected in quantitative attempts to explain racial disparities.

First, in decomposing the total racial health disparity in health ( $Y$ ), we do not consider any “pre-exposure” variables  $C$  (e.g., parental income) to be *confounders* of Race and health. Doing so implicitly treats Race as an individualized exposure that begins at birth, rather than part of a relational, multigenerational system of racism. Mediation analysis controlling for many “pre-exposure confounders” means decomposing a marginal racial disparity conditional on other variables treated as “not race,” such as parental material resources. While this construction of the causal model may be illuminating in explaining an artificial, hypothetical disparity, it is of limited utility in explaining the total observed disparity. Any racial disparity in health ( $Y$ ) is the result of historical and contemporary racism, and every observed “pre-exposure” variable that differs across racialized categories is more appropriately considered a *mediator* ( $M$ ) rather than a confounder.

Second, there are three distinct pathways through which household income ( $M$ ) might play a role in maintaining the association between self-identified race and health that are not separately identified in the Baron-Kenny approach (Jackson and VanderWeele 2019; Wang and Arah 2015), and each pathway maps onto a different facet of Reskin’s (2012) race discrimination system.

- **The controlled direct effect (CDE)**, Arrow (c). This includes all unobserved mediating pathways through which racism becomes embodied or otherwise influences  $Y$  that do not operate through the measured mediators – in other words, *unobserved racism*.
- **The proportion attributable to interaction (PAI)**, Arrow (a) + the interactive effect of  $\text{Race} * M$  through Arrow (b). This is a form of structural *racial*

*discrimination*, whereby the underlying system governing the relationship between  $M$  and  $Y$  operates differently across racialized categories.

- **The pure indirect effect (PIE)**, Arrow (a) + the main effect of  $M$  through Arrow (b). This is a form of *emergent discrimination*, whereby racism patterns the distribution of  $M$  via Arrow (a) but  $M$  affects  $Y$  regardless of racialized category.

As discussed above, in many studies the CDE is the target effect. The implicit logic is that this effect is somehow representing the “true effect of racial discrimination” because we have removed other indirect pathways connecting racial categories to  $Y$  which are conceptually bracketed as “not race” (e.g., income). Still, the CDE is completely dependent on which observed mediators are included in estimation and is therefore more accurately considered a measure of ignorance, or the proportion of the total racialized disparity in  $Y$  that cannot be predicted through observed mediating systems. Instead of focusing exclusively on parameterizing the CDE, researchers of racial health inequities should more broadly consider how a system of structural racism produced stratification across a particular cohort via  $M$ ; or, as described by Stewart (2008), “swimming upstream.”

Third, the conventional mediation analysis requires that Assumption 4 holds by assuming no variables in  $L$  are influenced by Race (no dotted arrow in Figure 1). This is theoretically untenable, as racism is implicated in patterning the distributions of virtually all other variables that we might consider confounding the relationship between household income ( $M$ ) and health ( $Y$ ) (Naimi et al. 2016). We might consider neighborhood poverty to be one of the confounders ( $L$ ) of the relationship between household income and health, as neighborhood context affects both one’s level of household income and health. However, it is well-established that racism produces variation in neighborhood context across racialized categories via residential segregation (Charles 2003; Massey and Denton 1993). In a life course perspective, then, neighborhood context and household income are also reciprocally intertwined over time (Sharkey and Elwert 2011; Wodtke, Harding, and Elwert 2011).

What does this do to the interpretation of how much of the racial disparity is “explained by” household income in the conventional mediation analysis above? Researchers are left with a tricky counterfactual: what would the racial disparity have been if  $M$  had been fixed at its reference category and all other variables ( $L$ ) had also been fixed at their reference categories? Studies often adjust for how neighborhood poverty ( $L$ ) confounds the effect of household income ( $M$ ) on health ( $Y$ ), but in doing so, over-control the mediating pathway through neighborhood poverty,  $\text{Race} \rightarrow L \rightarrow Y$ . In assessing the mediating effect of household income on the relationship between race and health, it is therefore impossible in the conventional regression framework above to both control for and not control for neighborhood poverty (Sharkey and Elwert 2011). This is part of the “all else equal” axiom: how much do those racialized as Black differ in  $Y$  compared to those racialized as white, *all else equal*? Regardless of intent and emphasis, this counterfactual frame focuses attention on the residual racial disparity (CDE) (Kohler-Hausmann 2019; Zuberi and Bonilla-Silva 2008). In attempts to control for confounding, conventional estimators are often describing marginal, artificial worlds that bear little resemblance to the reality of how the dynamic, relational *process* of racism operates over time. The quantitative field of racial health



disparities increasingly leans on the nuanced theoretical and conceptual models of race, racialization, and structural racism. Still, studies often rely on conventional estimators that restrict understanding of how racism operates over time and within and across domains to produce racialized disparities in health.

### New Developments in Quantitative Causal Mediation

In Figure 2, we consider three sets of mediators of the racial health disparity ( $M^{(1)}$ ,  $M^{(2)}$ ,  $M^{(3)}$ ) observed over three age windows. When decomposing a disparity in Health by self-identified race, we are describing a particular component of the historical causal process that connects the system of racism to population variation in health, following a particular racialized cohort through a particular space and time. Using self-identified race in this way provides an incomplete picture of the historical process of racialization and the effects of racism, which are malleable over space and time. But as many other authors have described in defending the need to monitor health indicators by race, measuring these disparities can still be useful in retrospectively identifying causal mechanisms maintaining and reproducing the racial stratification that leads to embodied health inequities for particular cohorts (Chowkwanyun and Reed 2020; Krieger 2018).

In this longitudinal framework, we are no longer conceptually separating variables into “race” and “not race” in the same way that might be implied by conventional regression estimators. For example, variation in  $M^{(2)}$  is explicitly defined as the historical product of several pathways:

1. Race is linked to the distribution of  $M^{(2)}$  through racism in unmeasured pathways not captured by  $M^{(1)}$  (*unobserved racism*).
2. Race is linked to the distribution of  $M^{(1)}$  through systems of racism.
  - a.  $M^{(1)}$  then influences  $M^{(2)}$  through an underlying system that racializes individuals and acts upon them differently (*racial discrimination*).
  - b.  $M^{(1)}$  then influences  $M^{(2)}$  regardless of racialized category (*emergent discrimination*).

The variables in  $M^{(2)}$  are thus not representing some proportion of variation in Health that is separable from Race (racism), but rather are explicitly part of the process of racism through which racial categories and Health become connected. Still, this decomposition is happening at a “high-level.” Each arrow is governed by its own underlying causal system, which includes specific actors producing and interacting with processes of racialization and racism (Ray 2019; Sewell 2016). For example, if we consider  $M^{(1)}$  to include neighborhood poverty, Sewell (2016) describes the political economy undergirding the Race  $\rightarrow M^{(1)}$  arrow, and Kramer et al. (2017) describe the processes of embodiment underlying the  $M^{(1)} \rightarrow Y$  arrow. Goosby et al. (2018) describe how stress-related biological mechanisms related to discrimination affect health outcomes, experiences which could be implicated in the causal arrow connecting Race  $\rightarrow Y$  in Figure 2 and within the systems of racial discrimination

governing interactive effects of mediators (e.g. racial discrimination influencing the effect of  $M^{(1)} \rightarrow Y$ ).

Our primary goal in this paper is not to specify every pathway by further differentiating each arrow in Figure 2. Instead, we aim to characterize the high-level structural system producing racialized disparities in an indicator of early adult health (cardio-metabolic risk), among this particular cohort, as well as the relative magnitudes of each of these broad mechanisms, in a way that reflects our theoretical framework. Though we use commonly available individual-level socioeconomic variables, such as educational attainment, we understand these to proxy structural relations in the systems of exchange between educational attainment and health outcomes (e.g., housing and labor markets).

### The Present Study

The present study uses causal mediation methods that are informed by a critical theory of racism as a fundamental cause of health inequities. The focus of this empirical case study is not on the process of racialization or the proximal mechanisms of interpersonal racism, but on the causally interconnected, reciprocal systems of structural racism (Bailey et al. 2017; Reskin 2012). We consider the process of a cohort aging through a system of racism as a causal inference problem of multiple mediators, where each intermediate exposure affects all others over time, in ways that vary by how an individual is racialized (Bailey et al. 2017; Esposito 2019). We argue that intentionally or not, the restrictive assumptions of conventional regression estimators for mediation analysis reify a problematic paradigm of race as a separable individual risk factor, rather than examining racism as a time-varying causal process (Zuberi and Bonilla-Silva 2008). We apply a g-formula method for multiple causal mediation to decompose the total racial disparity in cardio-metabolic risk in a longitudinal cohort of young adults. This decomposition consists of the following three effect pathways, which all represent different facets of racism: the controlled direct effect (“*unobserved racism*”), the portion attributable to interaction for each mediator (“*racial discrimination*”), and the pure indirect effect for each mediator (“*emergent discrimination*”). Under the assumption of no unmeasured confounding, we provide causal evidence for multiple mediating pathways through which structural racism produced embodied cardio-metabolic disparities for this cohort. We describe the relative magnitude of mediating effects through these systems, including neighborhood poverty, educational attainment, and household income.

## DATA AND METHODS

### Data

We use data from the National Longitudinal Study of Adolescent to Adult Health (Waves I, III, and IV: 1994–2008). Our outcome is a continuous measure of *cardio-metabolic risk* at Wave IV, defined as the first principal component of five biomarker variables indicating: 1) elevated waist circumference, 2) elevated blood pressure, 3) elevated triglycerides, 4) reduced high-density lipoprotein, and 5) pre-diabetic value of glycosylated hemoglobin (HbA1c). Previous research has demonstrated that this measure is a leading risk factor for

cardiovascular disease in the United States (Kane 2018). We normalize this index to have a mean of zero and variance of one.

We define each wave of data collection as “Adolescence” (Wave I; ages 12–18), “Transition to adulthood” (Wave III; ages 18–25), and “Young adulthood” (Wave IV; ages 25–32) and control for continuous years of age at the time of interview and self-identified sex in all models. Focusing on the role of socioeconomic factors, we consider the following causally ordered variables to mediate the relationship between self-identified race and cardio-metabolic risk in young adulthood (Figure 2):

- Adolescence ( $M^{(1)}$ ): annual household income (continuous US dollars), parental educational attainment (1 = at least one parent completing college), tract-level poverty rate (continuous).
- Transition to adulthood ( $M^{(2)}$ ): tract-level poverty rate (continuous).
- Young adulthood ( $M^{(3)}$ ): college attainment (1 = respondent completed college), tract-level poverty rate (continuous), annual household income (continuous US dollars).

In the final analysis, all mediators except neighborhood poverty rate are interacted with race (we include this interaction in online Appendix Figure 1). We exclude the interaction of neighborhood poverty rate in the final results because of concerns related to common support (i.e., lack of overlap in the distributions of neighborhood poverty between individuals racialized as Black and white). The Add Health cohort is reflective of the broader landscape of extreme Black-white residential segregation in that there is little overlap across this variable in our data. This can produce misleading conclusions based on interactions that are largely extrapolated outside of the observed data (Entwisle 2007; Oakes et al. 2015). Following Wodtke et al. (2011), our measure of neighborhood poverty rate is included in the analysis as cumulative over an individual’s entire life. We use the restricted Add Health sample, which includes 10,052 self-identified white individuals (25,888 person-years) and 3,893 self-identified Black individuals (9,741 person-years). We use multiple imputation to handle missing data (see online Appendix A for more information on analytic procedures).

### Parametric G-computation

The “g-formula” is a generalization of standardization that allows for the estimation of unconfounded summary effects without relying on the restrictive cross-world independence assumption. In conventional regression models (e.g. Baron-Kenny mediation) or demographic decomposition (e.g. Kitagawa decomposition), estimates of counterfactual change are calculated under the assumption that no other conditional probabilities change as a result of the exposure changing. In contrast, g-formula standardization makes explicit the sum of all “cascades” of conditional probabilities for all variables as the cohort ages through that time and space, consistent with Stewart’s (2008) “swimming upstream” (Figure 2). We use the g-formula, described in online Appendix A, to decompose the observed Black-white disparity in cardio-metabolic risk, using fully interacted models to account for the racialized processes underlying all systems in Figure 2.

We first present descriptive statistics for the outcome and all mediating variables by self-identified race. We then present the effect decomposition for the observed difference in cardio-metabolic risk scores between the populations racialized as Black and white. We map relevant research questions for each decomposed effect to their conceptual definitions in Table 1, using the specific example of educational attainment to demonstrate main ideas.

## RESULTS

Table 2 provides summary statistics by race. The reference category for all binary mediators is the value hypothesized to be associated with higher cardio-metabolic risk (e.g., 0 = no college attainment). We observe a significant disparity in our index of cardio-metabolic risk in young adulthood: 0.35 standard deviations higher in the cohort racialized as Black compared to white. We also observe large disparities across all mediating variables. The average neighborhood poverty rate for Black respondents in adolescence was 26 percent compared to just 11 percent for white respondents, a disparity which largely remained into adulthood. In addition, white adolescents lived in households with an average income of \$51,66 compared to \$30,496 for Black adolescents. A smaller percentage of Black adolescents lived in a household where a parent completed college compared to white adolescents (16 vs. 24 percent, respectively). In young adulthood, the disparity in college attainment was 22% vs. 33%, suggesting a general increase in college attainment across generations but also an increase in the Black-white gap.

Online Appendix Table A.1 displays parameter estimates from all mediator and outcome models. These are used to parameterize conditional probabilities across the life course in the g-formula and should not be given a causal interpretation. Table 3 provides results from the mediation analysis. In this counterfactual decomposition, we divide each decomposed effect by the total effect to determine the proportion of the racial health disparity that would be eliminated if that effect pathway had not operated on this cohort as they aged from adolescence to young adulthood (Table 1). For example, results in Table 3 reveal that if the cohort racialized as Black had instead been racialized as white by the system connecting educational attainment to cardio-metabolic risk, we would expect the total disparity in cardio-metabolic risk by adulthood to be reduced by 0.06 standard deviations (−0.11 to −0.01,  $p < 0.05$ ), or 18% of the total observed disparity. Figure 3 illustrates each decomposed effect.

Results indicate that roughly 58% of the total racial disparity in the index of adult cardio-metabolic risk can be explained by the observed mediating pathways (Table 3, Figure 3). Three of these mediating effect pathways are statistically significant ( $p < 0.05$ ): the pure indirect effect (PIE) operating via cumulative exposure to neighborhood poverty across the life course (−0.07; −0.12 to −0.02, 21%), the proportion attributable to interaction (PAI) of whether a parent of the respondent completed college (−0.05; −0.09 to −0.01, 15%), and the proportion attributable to interaction (PAI) of whether the respondent completed college (−0.06; −0.11 to −0.01, 18%). The remaining 42% of the observed racial disparity (CDE) is explained by unobserved mediating pathways operating outside the measures included in this analysis.

The two interactive effect pathways that are most important in explaining observed disparities in cardio-metabolic risk involve returns to parental educational attainment and returns to personal educational attainment. These pathways both serve as examples of structural racism via *racial discrimination* in the causal systems governing the returns to education on adult cardio-metabolic health in the United States during this particular period. Our results indicate that, in some way, that system connecting educational attainment to health treated individuals racialized as Black differently than those racialized as white.

In contrast, results indicate that cumulative exposure to neighborhood poverty is the most important pure indirect effect contributing to the racial disparity in cardio-metabolic risk (Table 3), which serves as an example of structural racism via *emergent discrimination* (−0.07; −0.12 to −0.02, 21%). Living in a neighborhood with high poverty impacts health regardless of racialized category, but the racialized distributions of this harmful exposure are highly unequal in ways that contribute to racial health inequality. In this cohort, 16 percent of respondents racialized as white resided in neighborhoods with a poverty rate exceeding 20 percent in adolescence, compared with 61 percent of those racialized as Black. In the g-formula framework, we are able to account for the fact that neighborhood poverty at one age influences all subsequent exposures (as shown in Figure 2). However, our estimate of this particular indirect effect (21 percent) is net of additional indirect pathways (the arrows directly connecting this neighborhood poverty to cardio-metabolic risk in Figure 2), suggesting the salience of neighborhood context across the life course in shaping adult health, beyond its cumulative impact on adulthood education and income.

## DISCUSSION

### Structural Racism as a Relational Process

If race, racialization, and racism are dynamic relational processes then our quantitative methods for studying racial inequality must move beyond conventional regression models, which largely treat “race” as a static trait of individuals that is separable from racialized constructs. In this study, we have provided a worked example of a life course mediation framework for decomposing racial disparities in adult cardio-metabolic function. We have clarified the relative importance of several mediators within the system of racism experienced by this particular cohort, such as the importance of *emergent discrimination* across the life course via disproportionate exposure to neighborhood poverty and *racial discrimination* via the system governing the relationship between educational attainment and adult health. In doing so, this study serves as one example by which a self-identified race variable can be better used to explain the causal process of racism in shaping outcomes in quantitative studies.

Using nationally representative longitudinal data, we decomposed the Black-white disparity in cardio-metabolic risk in young adulthood, paying particular attention to the roles of life course socioeconomic exposures in the production of the racialized health inequity. Our results revealed a sizable disparity in cardio-metabolic risk within this cohort. We decomposed this disparity using several socioeconomic pathways, implicating cumulative dynamics of structural racism arising from both racial discrimination and emergent discrimination across the life course. We understand the production of this health disparity

as a relational process in which various structural mediating systems create differential exposures based on racialized category (e.g., *emergent discrimination*) and treat individuals differently based on racialized category (e.g., *racial discrimination*).

A large proportion of the total disparity in cardio-metabolic risk – 58 percent – was explained by the observed mediators. Consistent with previous research (Boen 2016; Phelan and Link 2015), these results indicate the prominence of socioeconomic factors in the production of racialized health disparities. Much of this finding results from neoliberal political arrangements in the United States focused on the coupling of well-being to individualized “human capital,” as well as how this process interacts with contemporary and historical racism (Bailey et al. 2017; Laster Pirtle 2020; Muntaner et al. 2010) to produce racialized socioeconomic distributions. We further found that parental socioeconomic status shapes adult health, with racial disparities in parental SES playing a key role in the generation of racial health inequities. Parental SES shapes health through many mediators observed here across the life course (e.g., children’s eventual likelihood of college completion), but we demonstrate that they also have significant indirect effects.

Our results indicated that differential *returns* to parental educational attainment and personal educational attainment also produced racialized health disparities. Reskin (2012) highlights these as examples of structural racism via *racial discrimination* in the causal systems governing the returns to education on adult cardio-metabolic health. This finding, consistent with the “differential returns hypothesis,” (Boen 2016; Esposito 2019) could be due to many factors, including how systems of education convey advantage in racialized ways that may eventually impact health outcomes (e.g., occupational status, access to social and professional networks, stress, etc.). A key take-away from this finding is that equalizing access to education would be insufficient in closing Black-white health gaps without parallel efforts to dismantle racism in other domains.

Our results further indicated that exposure to neighborhood poverty contributed to the observed racial disparity in cardio-metabolic risk, which serves as an example of structural racism via what Reskin (2012) terms *emergent discrimination*. While Figure 2 shows how cumulative neighborhood poverty across the life course influenced all subsequent exposures, we also observed pure indirect effects of neighborhood poverty on adult cardio-metabolic risk. In these ways, results are consistent with research pointing to the prominence of childhood and adolescence as sensitive periods for health and development, when exposure to neighborhood poverty shapes future patterns of health directly and indirectly by shaping adult education, occupation, income, and neighborhood context (Kravitz-Wirtz 2016).

Roughly 42 percent of the Black-white gap in cardio-metabolic risk was explained by unobserved mediating pathways, including structural racism via sub-systems not accounted for in our analysis (e.g. punitive policing and court systems, access to healthcare, housing, etc.) and biopsychosocial racism-related stress pathways (Alexander 2012; Boen 2020; Goosby et al. 2018). To the extent these unmeasured systems operate along the causal pathways accounted for in our analysis, they will be included in those effects, but could be further decomposed. For example, contact with policing and differential healthcare access may mediate both the relationships between 1) racialized category and adult health (i.e.,



explain a portion of the CDE) and 2) the relationship between neighborhood poverty in adolescence and adult health (i.e., explain a portion of that PIE).

### Constructing Causal Models to Explain Health Disparities

Importantly, our results are ultimately predicated on our choices in how we draw our DAG — choices that can only be justified by our theoretical conceptions of each construct, their causal ordering, and their causal relationships. The empirical goal of the current analysis is to characterize the impacts of multiple dependent mediators related to structural racism at a high level in shaping racial health inequities, but in doing so we leave many causal arrows as “black boxes” with insufficiently deep explanation. Extensive empirical work, both quantitative and qualitative, exists for any one of these arrows. Still, the empirical approach employed in this study shows promise for identifying and further examining the mediating pathways whereby structural racism contributes to population health gaps and putting those pathways in a broader context relative to the total real-world disparity.

By combining a race interacted model with an analysis of mediating pathways (Diez Roux 2012), our analysis supports a theoretical framework grounded in the causal dependence between racism and class relations as a fundamental cause of health in the United States (Laster Pirtle 2020; Phelan and Link 2015; Robinson and Bailey 2020). Our results speak to the critical roles of the unequal racialized distributions of socioeconomic risks, opportunities, and exposures—as well as racialized effects of these exposures—in the production of racial health disparities. Still, as Williams (2019: 4) notes:

“It may be more feasible for scholars to move away from asking ‘what factors account for racial inequality’ to asking, ‘what factors maintain racial inequality.’ The former question tends to lead to human capital (i.e., individualistic) explanations whereas the latter question lends itself to racism-based (i.e., structural) explanations.”

The causal arrows and their racialized interactions in Figure 2 do not arise naturally or inevitably, as is often implied by uncritical “social determinants of health” frameworks that do not consider relational systems of social and economic production. Instead, each causal arrow is maintained by a political economy of specific actors who stand to benefit from particular arrangements rooted in capitalist interests and white supremacy (Laster Pirtle 2020; Ray 2019; Roberts 2011; Sewell 2016).

### Scope and Limitations

Returning to the scope of our study, we have described how the common statement that “*X% of the racial disparity in outcome Y is explained by M (‘non-race’ variable)*” is predicated on unlikely assumptions regarding post-treatment bias and inconsistent with social constructivist perspectives of the relation of racism and race. Still, in providing a worked example of a quantitative technique that can overcome these issues, we retain three important limitations that scholars should consider in future work: 1) we use readily available secondary data with limited information, 2) we use “self-identified race” to proxy the many fluid forms of Black-white racialization and historical and contemporary manifestations of racism underlying our broad model (see Sen & Wasow 2016), and 3) we

use individual-level measurements of social characteristics and material conditions to proxy more dynamic mediating systems of racialized social relations. Indeed, the distribution of income, for example, is the consequence of relational processes such as labor exploitation, employer discrimination, and wage theft; income in turn influences health through relations to largely privatized healthcare, housing, and more (see Muntaner et al. 2010, Pirtle Laster 2020, and Sewell 2016). Still, this represents what has been the modal setting of racial health disparities research: combining a limited self-identified race variable from a large national cohort study (e.g., Add Health, Health and Retirement Study, Panel Study on Income Dynamics, etc.) with individual- and household-level indicators of socioeconomic position. As such, a goal of this paper is to provide an alternative counterfactual approach to conventional regression models that better aligns relational and social constructivist theories of race and racism with quantitative studies of racialized health disparities using longitudinal survey data. Our study is not prescriptive from a policy standpoint, but rather provides the scaffolding of a quantitative approach for how racism and racialization ultimately produce disparities in population health for a *particular cohort* aging through a *particular period and context of racism and racialization*, while avoiding common pitfalls of conventional regression analysis.

Each relation in our broad model can be further clarified via more specific information on the historic and contemporary racist relations connecting individuals to structures and the salient forms of racialization deployed. This work will involve the collection of new forms of data on racialized *relations* rather than static social determinants of health (Krieger 2018; Muntaner 2013; Muntaner et al. 2010; Prins et al. 2021); the merging of new data to cohort studies, such as contextual information on segregation (Sharkey 2013); reflexive qualitative inquiry (Ford and Airhihenbuwa 2010); description of historical context (Chowkwanyun 2011; Chowkwanyun and Reed 2020; D. Williams 2019); engagement with critical, intersectional, feminist, and postcolonial theories (Crenshaw 2017; Delgado and Stefania 2017; Glenn 2015); and an ecosocial framework describing the actual processes of embodiment (Krieger 2012). Importantly, quantitative attempts to parameterize the links between particular racist relations should be mindful of not losing sight of the “forest” of structural racism through the “trees” of any particular form of racist relations (Diez Roux 2011; Robinson and Bailey 2020). Alternatively, efforts to clarify the links between structural racism and health should be cautious of boiling down what is a complex, flexible structural arrangement to a single parameter (Diez Roux 2012).

There are other limitations to our study. The process of racism is gendered, among many other intersections, including (but not limited) to colorism, heterosexism, ageism, nativism, ableism, and economic exploitation (Crenshaw 1991; Ford and Airhihenbuwa 2010). This may (in its most simplified form) involve three-way interactions of all observed mediators with race and gender (Bauer and Scheim 2019), for example. By not including this additional interaction, we average over potential gendered heterogeneity.

In decomposing a population health disparity along the lines of self-identified race, we are focused on retrospectively examining the forward process by which racialized individuals moved through a system of racism (Stewart 2008). It is beyond the scope of this analysis to

also consider how these processes then serve to reify racial categories [for this, see Sewell (2016)].

Last, it is important to qualify the fundamental limitations of quantitative counterfactuals (Schwartz, Gatto, and Campbell 2011). We consider the counterfactual “*What if those racialized as Black had instead been racialized as white within the racist structural systems characterizing the time and place through which the Add Health cohort aged?*” to be based on a “well-described exposure” (e.g., racialization in a racist system) rather than necessarily a “well-defined intervention” (Glymour 2006; Hernán and Robins 2019; Holland 2008; Morgan and Winship 2014; Pearl 2014). We are not interpreting any of the decomposed effects in isolation as *what would actually happen* if we were to somehow equalize Black-white exposures or effects of exposures. But we agree with arguments made by Robinson & Bailey (2020) that precise quantitative identification of such hypothetical future causal effects is not a prerequisite for supporting a broad policy agenda aimed at dismantling pathways of structural racism affecting health, an agenda based on retrospective quantitative and qualitative causal triangulation, a robust theory of racism, and longstanding social movements (Bailey et al. 2017; Darity and Mullen 2020; Ford and Airhihenbuwa 2010; Taylor 2016).

## Conclusions

This study provided a worked example of a life course mediation framework for decomposing racial health disparities, but this framework can be generalized to any population disparity. We clarified the relative importance of several mediators within the system of racism experienced by the Add Health cohort, such as the importance of *emergent discrimination* across the life course via disproportionate exposure to neighborhood poverty experienced and *racial discrimination* via the system governing the relationship between educational attainment and adult health. However, substantial assumptions and limitations of this analysis (and any quantitative analysis of racial disparities) underscores the crucial importance of triangulation (Du Bois 1898; Itzigsohn and Brown 2020). Quantitative inference, predicated on estimating unobserved counterfactuals from observed data under strong assumptions of full or partial exchangeability, is only one mode of causal inference, not *the definition* of causal inference (Schwartz, Gatto, and Campbell 2017).

Much of modern quantitative causal inference in the study of health pushes research questions towards “well-defined interventions” seen as more “proximal” to the individual, which reinforces a neoliberal paradigm of social change predicated on marginal interventions that do not threaten structural or institutional arrangements (Robinson and Bailey 2020; Schwartz et al. 2016). At the same time, conventional regression models are subject to time-varying confounding issues, which can reify notions of individual race as a separable “risk factor” from all other social relations. But an uncritical application of more complex methods risks losing sight of the fundamental causes that govern the distribution of more proximal risk factors, and ultimately the total disparity in population health (Bailey et al. 2017; Jackson and Arah 2019; Krieger 2014; Phelan and Link 2015; Schwartz et al. 2017). It is therefore our hope that this study can offer one framework for leveraging developments in quantitative causal inference to decompose a population disparity observed

at a given time to the historical life course process through which individuals come to embody racist structural systems, supporting more holistic causal narratives rather than isolating effects of marginal changes.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS & FUNDING

We thank Irma Elo, Xi Song, Tukufu Zuberi, and Daniel Aldana Cohen for their helpful feedback on earlier drafts of this manuscript. We also thank Bridget Goosby and participants of the *Expanding Diversity of Biosocial Research: Opportunities and Challenges* session at the 2020 meeting of the American Sociological Association. This research uses data from the National Longitudinal Study of Adolescent to Adult Health (Add Health), a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill and funded by Grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>).

NG was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development Training Grant (T32-HD-007242–36A1). CB is grateful to the Population Studies Center at the University of Pennsylvania (National Institutes of Health’s Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH Grant No. R24 HD044964) and the Axilrod Faculty Fellowship program at the University of Pennsylvania for general support. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## Biographies

**Nick Graetz** is a postdoctoral research associate in the Department of Sociology at Princeton University. His work examines the intersection of population health, structural racism, and housing policy.

**Courtney E. Boen** is an assistant professor in the Department of Sociology at the University of Pennsylvania, whose research examines the structural and social determinants of population health inequality.

**Michael H. Esposito** is an assistant professor in the Department of Sociology at Washington University in St. Louis. His work examines the structural drivers of population health inequities.

## REFERENCES

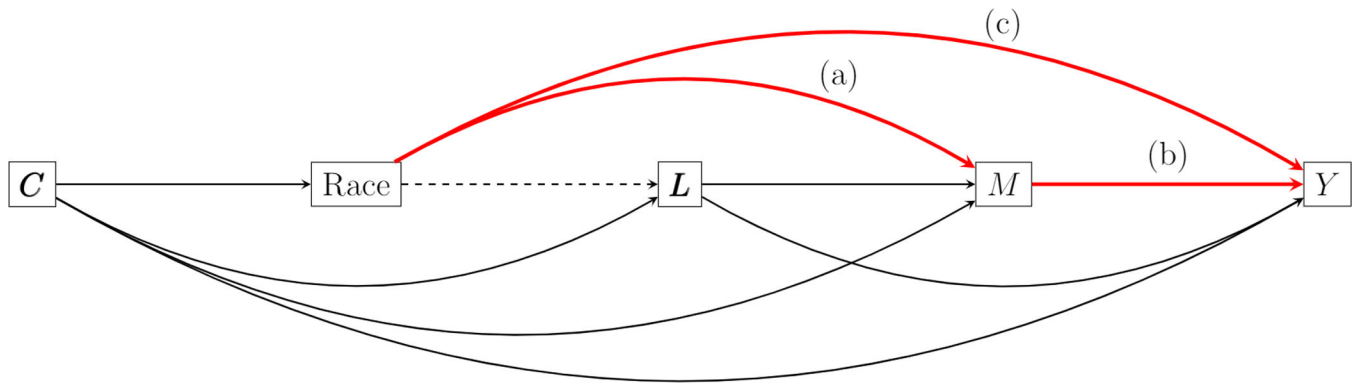
- Alexander Michelle. 2012. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. The New Press.
- Bailey Zinzi D., Feldman Justin M., and Bassett Mary T. 2020. “How Structural Racism Works — Racist Policies as a Root Cause of U.S. Racial Health Inequities.” *New England Journal of Medicine*.
- Bailey Zinzi D., Krieger Nancy, Agénor Madina, Graves Jasmine, Linos Natalia, and Bassett Mary T. 2017. “Structural Racism and Health Inequities in the USA: Evidence and Interventions.” *The Lancet* 389(10077):1453–63.
- Bauer Greta R., and Scheim Ayden I. 2019. “Methods for Analytic Intercategorical Intersectionality in Quantitative Research: Discrimination as a Mediator of Health Inequalities.” *Social Science and Medicine* 226:236–45. [PubMed: 30674435]

- Boen Courtney. 2016. "The Role of Socioeconomic Factors in Black-White Health Inequities across the Life Course: Point-in-Time Measures, Long-Term Exposures, and Differential Health Returns." *Social Science and Medicine* 170:63–76. [PubMed: 27764654]
- Boen Courtney. 2020. "Death by a Thousand Cuts: Stress Exposure and Black-White Disparities in Physiological Functioning in Late Life." *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences* 75(9):1937–50. [PubMed: 31127845]
- Du Bois WEB 1898. "The Study of the Negro Problems." *Annals of the American Academy of Political and Social Science*.
- Du Bois WEB 1899. *The Philadelphia Negro: A Social Study*. The University of Pennsylvania Press.
- Bonilla-Silva Eduardo. 1997. "Rethinking Racism: Toward a Structural Interpretation." *American Sociological Review* 62(3):465–80.
- Bonilla-Silva Eduardo. 2009. *Racism without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America*. Rowman & Littlefield Publishers.
- Brown Tyson H. 2018. "Racial Stratification, Immigration, and Health Inequality: A Life Course-Intersectional Approach." *Social Forces* 96(4):1507–40.
- Brown Tyson H., Richardson Liana J., Hargrove Taylor W., and Thomas Courtney S. 2016. "Using Multiple-Hierarchy Stratification and Life Course Approaches to Understand Health Inequalities: The Intersecting Consequences of Race, Gender, SES, and Age." *Journal of Health and Social Behavior* 57(2):200–222. [PubMed: 27284076]
- Charles Camille Zubrinsky. 2003. "The Dynamics of Racial Residential Segregation." *Annual Review of Sociology* 29(1):167–207.
- Chowkwanyun Merlin, and Reed Adolph L. 2020. "Racial Health Disparities and Covid-19 — Caution and Context." *New England Journal of Medicine* 383(3):201–3. [PubMed: 32374952]
- Cogburn Courtney D. 2019. "Culture, Race, and Health: Implications for Racial Inequities and Population Health." *The Milbank Quarterly* 97(3):736–61. [PubMed: 31512293]
- Crenshaw Kimberle. 1991. "Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color." *Stanford Law Review* 43(6):1241–99.
- Crenshaw Kimberle. 2017. *On Intersectionality: Essential Writings*. New York, NY: The New Press.
- Darity William A. Jr., and Kirsten Mullen A 2020. *From Here to Equality*. The University of North Carolina Press.
- Delgado Richard, and Stefanic Jean. 2017. *Critical Race Theory (Third Edition): An Introduction*. New York, NY: NYU Press.
- Diez Roux Ana V. 2011. "Complex Systems Thinking and Current Impasses in Health Disparities Research." *American Journal of Public Health* 101(9):1627–34. [PubMed: 21778505]
- Diez Roux Ana V. 2012. "Conceptual Approaches to the Study of Health Disparities." *Annual Review of Public Health* 33(1):41–58.
- Entwisle Barbara. 2007. "Putting People into Place." *Demography* 44(4):687–703. [PubMed: 18232206]
- Esposito Michael H. 2019. "Inequality in Process: Income and Heterogeneous Educational Health Gradients Among Blacks and Whites in the USA." *Race and Social Problems* (0123456789).
- Ford Chandra L., and Airhihenbuwa Collins O. 2010. "The Public Health Critical Race Methodology: Praxis for Antiracism Research." *Social Science and Medicine* 71(8):1390–98. [PubMed: 20822840]
- Frerichs Leah, Lich Kristen Hassmiller, Dave Gaurav, and Corbie-Smith Giselle. 2016. "Integrating Systems Science and Community-Based Participatory Research to Achieve Health Equity." *American Journal of Public Health* 106(2):215–22. [PubMed: 26691110]
- Geronimus Arline T., Hicken Margaret, Keene Danya, and Bound John. 2006. "'Weathering' and Age Patterns of Allostatic Load Scores among Blacks and Whites in the United States." *American Journal of Public Health* 96(5):826–33. [PubMed: 16380565]
- Glenn Evelyn Nakano. 2015. "Settler Colonialism as Structure." *Sociology of Race and Ethnicity* 1(1):52–72.

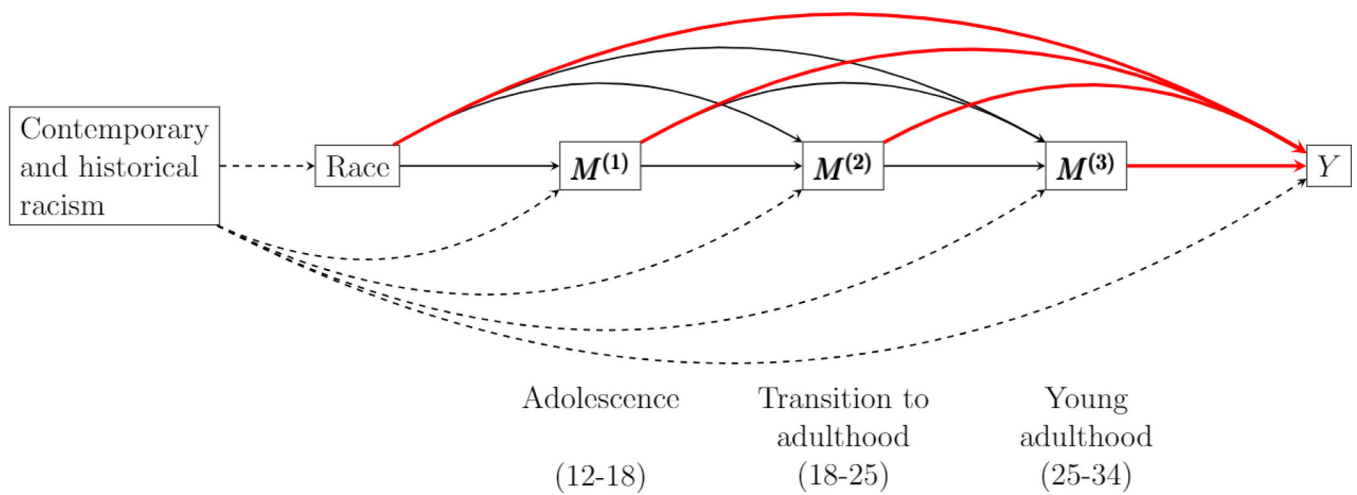
- Goosby Bridget J., Cheadle Jacob E., and Mitchell Colter. 2018. "Stress-Related Biosocial Mechanisms of Discrimination and African American Health Inequities." *Annual Review of Sociology* 44(1):319–40.
- Green Tiffany L., and Darity William A. 2010. "Under the Skin: Using Theories from Biology and the Social Sciences to Explore the Mechanisms behind the Black-White Health Gap." *American Journal of Public Health* 100(SUPPL. 1):S36. [PubMed: 20147678]
- Holland P 2008. "Causation and Race." Pp. 93–109 in *White Logic, White Methods*, edited by Zuberi T and Bonilla-Silva E Rowman & Littlefield Publishers.
- Itzigsohn Jose, and Brown Karida L. 2020. "The Sociology of W. E. B. Du Bois: Racialized Modernity and the Global Color Line." NYU Press. Retrieved March 19, 2021 (<https://nyupress.org/9781479856770/the-sociology-of-w-e-b-du-bois/>).
- Jackson John W., and Arah Onyebuchi A. 2019. "Invited Commentary: Making Causal Inference More Social and (Social) Epidemiology More Causal." *American Journal of Epidemiology* 189(3):179–82.
- Jackson John W., and VanderWeele Tyler J. 2019. "Intersectional Decomposition Analysis with Differential Exposure, Effects, and Construct." *Social Science and Medicine* 226(January):254–59. [PubMed: 30770131]
- Kohler-Hausmann Issa. 2019. "Eddie Murphy and the Dangers of Counterfactual Causal Thinking about Detecting Racial Discrimination." *Northwestern University Law Review* 113(5):1163–1228.
- Kramer Michael R., Schneider Eric B., Kane Jennifer B., Margerison-Zilko Claire, Jones-Smith Jessica, King Katherine, Davis-Kean Pamela, and Grzywacz Joseph G. 2017. "Getting Under the Skin: Children's Health Disparities as Embodiment of Social Class." *Population Research and Policy Review* 36(5):671–97. [PubMed: 29398742]
- Kravitz-Wirtz Nicole. 2016. "Cumulative Effects of Growing Up in Separate and Unequal Neighborhoods on Racial Disparities in Self-Rated Health in Early Adulthood." *Journal of Health and Social Behavior* 57(4):453–70. [PubMed: 27799591]
- Krieger Nancy. 2014. "On the Causal Interpretation of Race." *Epidemiology* 25(6):937.
- Krieger Nancy. 2018. "The Science and Epidemiology of Racism and Health: Racial/Ethnic Categories, Biological Expressions of Racism, and the Embodiment of Inequality — an Ecosocial Perspective." *What's the Use of Race?* (2010).
- Laster Pirtle Whitney N. 2020. "Racial Capitalism: A Fundamental Cause of Novel Coronavirus (COVID-19) Pandemic Inequities in the United States." *Health Education and Behavior* 47(4).
- Lundberg Ian, Johnson Rebecca, and Stewart Brandon. 2021. "What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory." *American Sociological Review* Forthcomin.
- Massey Douglas S., and Denton Nancy A. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Harvard University Press.
- Muntaner Carles. 2013. "Invited Commentary: On the Future of Social Epidemiology--a Case for Scientific Realism." *American Journal of Epidemiology* 178(6):852–57. [PubMed: 24008904]
- Muntaner Carles, Borrell Carme, Vanroelen Christophe, Chung Haejoo, Benach Joan, Kim Il Ho, and Ng Edwin. 2010. "Employment Relations, Social Class and Health: A Review and Analysis of Conceptual and Measurement Alternatives." *Social Science and Medicine* 71(12):2130–40. [PubMed: 21075495]
- Naimi Ashley I., Schnitzer Mireille E., Moodie Erica E. M., and Bodnar Lisa M. 2016. "Mediation Analysis for Health Disparities Research." *American Journal of Epidemiology* 184(4):315–24. [PubMed: 27489089]
- Oakes J. Michael. 2009. "Commentary: Individual, Ecological and Multilevel Fallacies." *International Journal of Epidemiology* 38(2):361–68. [PubMed: 19174541]
- Oakes J. Michael, Andrade Kate E., Biyoow Ifrah M., and Cowan Logan T. 2015. "Twenty Years of Neighborhood Effect Research: An Assessment." *Current Epidemiology Reports* 2(1):80–87. [PubMed: 28133588]
- Phelan Jo C., and Link Bruce G. 2015. "Is Racism a Fundamental Cause of Inequalities in Health?" *Annual Review of Sociology* 41(1):311–30.



- Ray Victor. 2019. "A Theory of Racialized Organizations." *American Sociological Review* 84(1):26–53.
- Reskin Barbara. 2012. "The Race Discrimination System." *Annual Review of Sociology* 38(1):17–35.
- Roberts Dorothy E. 2011. *Fatal Invention: How Science, Politics, and Big Business Re-Crete Race in the Twenty-First Century*. New Press.
- Robinson Whitney R., and Bailey Zinzi D. 2020. "Invited Commentary: What Social Epidemiology Brings to the Table - Reconciling Social Epidemiology and Causal Inference." *American Journal of Epidemiology* 189(3):171–74. [PubMed: 31566211]
- Schwartz Sharon, Gatto Nicolle M., and Campbell Ulka B. 2011. "What Would Have Been Is Not What Would Be: Counterfactuals of the Past and Potential Outcomes of the Future." *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures* 25–46.
- Schwartz Sharon, Gatto Nicolle M., and Campbell Ulka B. 2017. "Heeding the Call for Less Casual Causal Inferences: The Utility of Realized (Quantitative) Causal Effects." *Annals of Epidemiology* 27(6):402–5. [PubMed: 28641760]
- Schwartz Sharon, Prins Seth, Campbell Ulka, and Gatto Nicolle. 2016. "Is the 'Well-Defined Intervention Assumption' Politically Conservative?" *Social Science & Medicine* 166:254–57. [PubMed: 26777446]
- Sen Maya, and Wasow Omar. 2016. "Race as a Bundle of Sticks: Designs That Estimate Effects of Seemingly Immutable Characteristics." *Annual Review of Political Science* 19:499–522.
- Sewell Abigail A. 2016. "The Racism-Race Reification Process." *Sociology of Race and Ethnicity* 2(4):402–32.
- Sharkey Patrick, and Elwert Felix. 2011. "The Legacy of Disadvantage: Multigenerational Neighborhood Effects on Cognitive Ability." *American Journal of Sociology* 116(6):1934–81.
- Stewart Quincy Thomas. 2008. "Swimming Upstream: Theory and Methodology in Race Research." in *White Logic, White Methods*.
- Taylor Keeanga-Yamahatta. 2016. *From #BlackLivesMatter to Black Liberation*. Haymarket Books.
- Wang Aolin, and Arah Onyebuchi A. 2015. "G-Computation Demonstration in Causal Mediation Analysis." *European Journal of Epidemiology* 30(10):1119–27. [PubMed: 26537707]
- Williams David R. 2012. "Miles to Go before We Sleep: Racial Inequities in Health." *Journal of Health and Social Behavior* 53(3):279–95. [PubMed: 22940811]
- Williams David R., Priest Naomi, and Anderson Norman B. 2016. "Understanding Associations among Race, Socioeconomic Status, and Health: Patterns and Prospects." *Health Psychology* 35(4):407–11. [PubMed: 27018733]
- Williams Deadric T. 2019. "A Call to Focus on Racial Domination and Oppression: A Response to "Racial and Ethnic Inequality in Poverty and Affluence, 1959–2015"." *Population Research and Policy Review* 38(5):655–63.
- Wodtke Geoffrey T., Harding David J., and Elwert Felix. 2011. "Neighborhood Effects in Temporal Perspective: The Impact of Long-Term Exposure to Concentrated Disadvantage on High School Graduation." *American Sociological Review* 76(5):713–36. [PubMed: 22879678]
- Zuberi Tukufu. 2001. *Thicker than Blood: How Racial Statistics Lie*. University of Minnesota Press.
- Zuberi Tukufu., and Bonilla-Silva Eduardo. 2008. *White Logic, White Methods: Racism and Methodology*. Rowman & Littlefield Publishers.
- Zuberi Tukufu, Patterson Evelyn J., and Stewart Quincy Thomas. 2015. "Race, Methodology, and Social Construction in the Genomic Era." *Annals of the American Academy of Political and Social Science* 661(1):109–27.

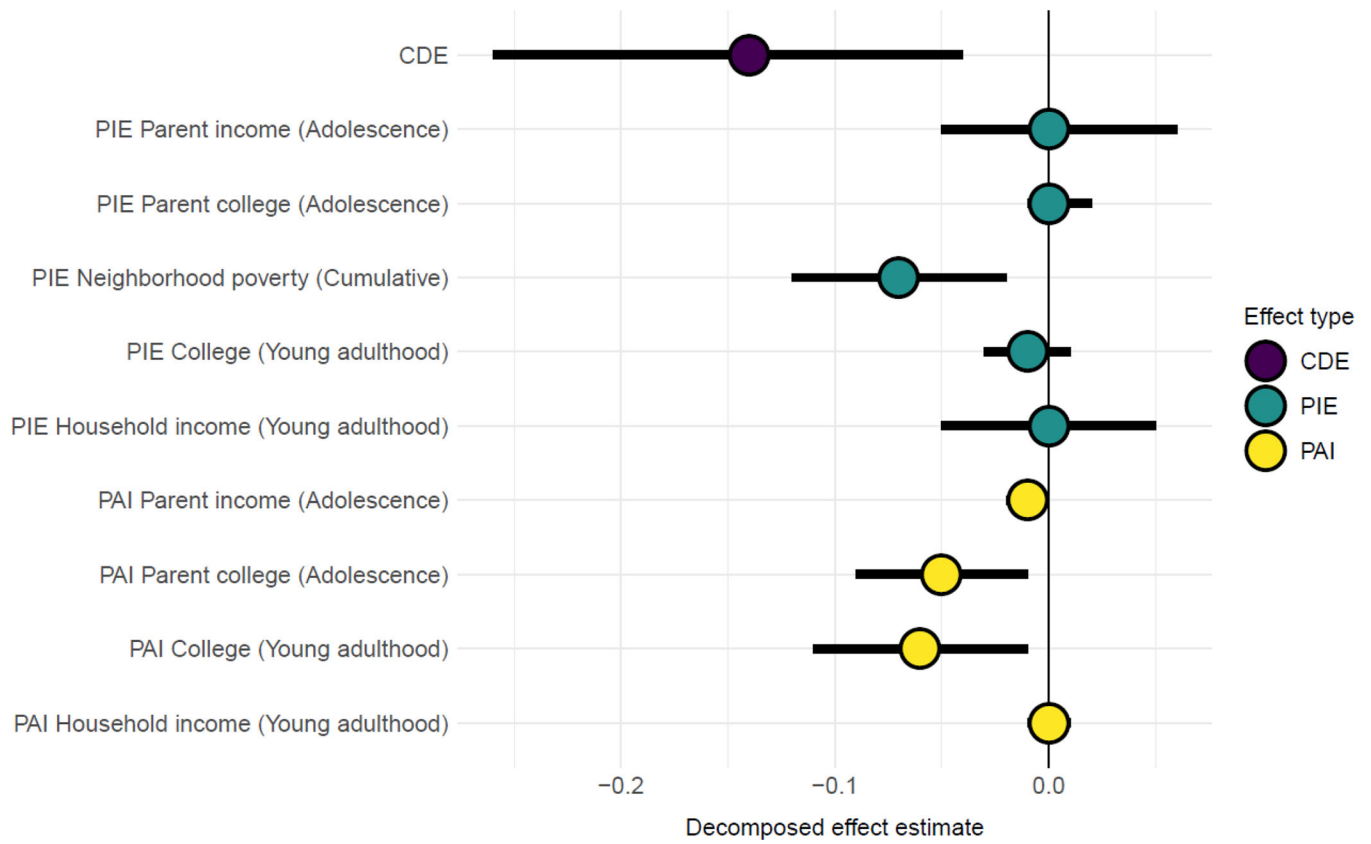


**Figure 1.** Directed-acyclic graph describing how the connection between Race and Health ( $Y$ ) is mediated by material deprivation, denoted  $M$ . Includes confounders of  $M \rightarrow Y$  denoted  $L$ , and Arrow (b) includes an interaction between Race and  $M$  in influencing  $Y$ .



**Figure 2.** Directed-acyclic graph describing the relationship between race and cardio-metabolic risk by adulthood.

**Note:** For visual clarity, two pieces of information are suppressed here compared to Figure 1: 1) Multiple observed mediators measured at the same time point are suppressed to a single node and set of arrows (e.g.  $M^{(1)}$ ) and 2) all arrows except those originating from Race contain an interaction with Race.



**Figure 3.**

Decomposition estimates for the total Black-white disparity in cardio-metabolic risk

**Notes:** The x-axis indicates the counterfactual change in the normalized cardio-metabolic risk index (i.e. standard deviations) attributable to that pathway had the population racialized as Black instead been treated as the population racialized as white by all mediating systems.

**Table 1.**  
Decomposed effect estimates and interpretation (adapted from Wang & Arah 2015).

Effect	Research question / counterfactual interpretation
ATE (total disparity)	What if those racialized as Black had been racialized and treated as white by all measured and unmeasured mediating systems?
CDE (disparity without mediators)	What if those racialized as Black had instead been racialized and treated as white in the absence of all measured mediators (i.e., how did <i>unobserved</i> mediating pathways of racism produce racialized disparities?)
PAI (disparity attributable to mediating interaction, e.g. via college attainment)	What if those racialized as Black had instead been racialized and treated as white by the underlying causal system (e.g., how did the system connecting educational attainment to health racialize individuals and discriminate based on those racial categories, via the <i>interaction</i> effect)?
PIE (disparity attributable to mediating main effect, e.g. via college attainment)	What if those racialized as Black had instead been racialized as white by the system producing the distribution of exposures but still racialized as Black by the system connecting the exposure to health (e.g., how much of the impact of education in mediating the total racial disparity was because of racialized differences in attainment rates and a <i>main</i> effect of attainment on health, via <i>emergent discrimination</i> )?

**Table 2.**

Descriptive statistics.

	White		Black		<i>p</i> -value
	Mean or prop.	SE	Mean or prop.	SE	
<b>Adolescence (ages 12–18)</b>					
Age (years)	15.90	(.02)	16.18	(.04)	***
Sex (1 = male)	.49	(.01)	.50	(.01)	***
Parent college attainment (1 = at least one parent completed college)	.24	(.01)	.16	(.01)	***
Parent household income (US\$)	\$51,666	(474)	\$30,496	(811)	***
Tract-level poverty rate (cont.)	.11	(.00)	.26	(.00)	***
<b>Transition to adulthood (ages 18–25)</b>					
Age (years)	22.09	(.02)	22.37	(.04)	***
Tract-level poverty rate (cont.)	.13	(.00)	.23	(.00)	***
<b>Young adulthood (ages 25–32)</b>					
Age (age)	28.92	(.02)	29.22	(.04)	***
Tract-level poverty rate (cont.)	.13	(.00)	.23	(.00)	***
College attainment (1 = respondent completed college)	.33	(.01)	.22	(.01)	***
Household income (US\$)	\$63,648	(519)	\$43,246	(844)	***
Cardio-metabolic risk (cont.)	-.08	(.02)	.27	(.03)	***
Person-years	25,888		9,741		
Unique individuals	10,052		3,893		

Weighted estimates. *p*-value of two-sided t-test:\* *p* < 0.05;\*\* *p* < 0.01;\*\*\* *p* < 0.001



**Table 3.**

G-formula effect decomposition.

Effect	Mean	Confidence interval	% of total disparity
ATE (total disparity)	-.33	(-.39, -.26)***	100
CDE (disparity without mediators)	-.14	(-.26, -.04)**	42
PIE (Parent income, Adolescence)	.00	(-.05, .06)	0
PIE (Parent college attainment, Adolescence)	.00	(-.01, .02)	0
PIE (Cumulative tract-level poverty)	-.07	(-.12, -.02)***	21
PIE (College attainment, Young adulthood)	-.01	(-.03, .01)	3
PIE (Household income, Young adulthood)	.00	(-.05, .05)	0
PAI (Parent income, Adolescence)	-.01	(-.02, .00)	3
PAI (Parent college attainment, Adolescence)	-.05	(-.09, -.01)*	15
PAI (College attainment, Young adulthood)	-.06	(-.11, -.01)*	18
PAI (Household income, Young adulthood)	.00	(-.01, .01)	0

\* p < 0.05;

\*\* p < 0.01;

\*\*\* p < 0.001