# RETHINKING RACE AND HEALTH: CAUSAL INFERENCE, INSTRUMENTATION AND EMPIRICAL MODELS

**by**

**GEOFF B. DOUGHERTY, MPH**

**A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy**

**Baltimore, Maryland**

**April, 2018**

## Abstract

Compared to whites in the United States, blacks experience heightened risk of many diseases as well as worse outcomes after medical treatment. Those differences are a major focus of epidemiologic investigation. Research has been complicated by two related issues: the absence of a precise definition of the biologic and social dynamics represented by the race variable, and the purportedly immutable nature of race, which hinders application of the potential outcomes framework to questions of race and health. We propose a conceptual framework separating race into distinct, modifiable components, including area-level structural racism and physical phenotype. These two components form a gene-environment interaction in which the effect of structural racism varies across phenotypic category. We identify structural racism as a key determinant of health, develop and validate a measure of county-level structural racism that includes multiple items representing differential institutional treatment of blacks and whites. We use a factor analysis model that accounts for measurement error and the correlated nature of potential indicators of structural racism. Our study addresses gaps in the content validity of previous work by measuring racism with indicators in five domains: employment, education, housing, healthcare and criminal justice. We estimate a structural racism factor score for counties representing 92% of the U.S. population. The model has adequate fit and strong construct validity. Finally, we evaluate the association between structural racism and BMI using data from the Behavioral Risk Factor Surveillance System. Following the relationships outlined in the conceptual framework, we specify an interaction between county structural racism sex, and black race on BMI. Predicted BMI for white males was substantively unchanged across levels of CSR, while BMI for white

women fell 0.4 kg/m$^{2\ with}$ a change in CSR from -1SD to 1 SD. For black males, a similar change in CSR resulted in a BMI increase of 0.4 kg/m$^2$. Black females reported the highest BMI across all levels of CSR, with BMI rising 0.1 kg/m$^2$ as CSR increased from -1 SD to 1 SD. The interaction term was statistically significant ($p<0.05$).

## Advisers and readers

**Lorraine T. Dean, ScD (Academic Co-Advisor)**
Assistant Professor
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health

**Bryan Lau, PhD (Academic Co-Advisor)**
Associate Professor
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health

**Stefanie DeLuca, PhD (Thesis Reader)**
Professor
Department of Sociology
Johns Hopkins University

**Elizabeth Colantuoni, PhD (Thesis Reader)**
Associate Scientist
Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health

**Renan Castillo, PhD (Alternate Thesis Reader)**
Associate Professor
Department of Health Policy and Management
Johns Hopkins Bloomberg School of Public Health

**Alden Gross, PhD, MHS (Alternate Thesis Reader)**
Assistant Professor
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health

## Acknowledgements

It has been my good fortune to work with three passionate and talented advisors. Dr. Thomas Glass encouraged me to grapple with a complicated and hazily framed research question, and underscored the centrality of theory and measurement to the scientific enterprise. Drs. Lorraine Dean and Bryan Lau gracefully agreed to take on me, and this project, midstream. They have helped me to frame, test, and present the hypotheses underlying this dissertation, and have done so with empathy and humor, as well as keen intellect.

The other members of the dissertation committee, Drs. Moyses Szklo, Sherita Golden, Stefanie DeLuca, and Justin Lessler, have provided valuable counsel throughout. Dr. Elizabeth Colantuoni served as a reader. Her advice on resolving the thorny issues around survey weighting and multilevel modeling were crucial to completing Aim 3. Dr. Alden Gross served as an alternate reader and key coauthor for Aim 2, advising on multiple rounds of confirmatory factor modeling and construct validity assessments. My fellow travelers in social epidemiology, including Usama Bilal, Emily Knapp, and Michael Windle, are a testament to the value of scientific community and collaboration.

My wife, Sara, agreed to uproot our life in Chicago and move to Baltimore so that I could become a (then) 44-year-old graduate student. That was a nice thing for her to do. My son, Sebastian, has been a kind and joyful companion on this voyage, even when I've missed day-care pickups and visits to the park to work on the dissertation. At last, I have an answer for Sebastian's frequently asked question: "When will you be done dissertating?"

# Table of contents

## List of tables

## List of figures

# Introduction

Compared to whites in the United States, blacks experience heightened risk of many diseases, as well as worse outcomes after medical treatment. Studies aimed at understanding the cause of these disparities, and evaluating the effect of interventions, are a major focus of public health researchers.

Efforts to understand and ameliorate disparities, though, have been complicated by two related issues. When researchers report racial disparities in disease incidence and treatment outcomes, it is unclear what the "race effect" actually signifies. The race variable has, over the past several decades, been interpreted variously as a visible manifestation of a genotype associated with poor health, a marker of exposure to cultural and behavioral patterns that cause ill health, and a proxy for exposure to structural and interpersonal racism. The lack of a precise and widely accepted accounting of the social and biologic dynamics represented by the race effect is an obstacle to understanding the role of race in etiology, and in the pursuit of potential interventions.

Further complicating this picture is the question of causal inference. The potential outcomes framework, which has become the primary means by which etiologic inference is made, specifies that a causal interpretation to study findings may only emerge from a design featuring a plausible intervention. Thus, the question remains as to whether one can estimate the causal effect of race on a health outcome, given that race is typically deemed an immutable trait.

The first aim of this dissertation is development of a conceptual framework that decomposes the "race effect" into distinct components, allowing researchers to assess which of those components are modifiable and thus amenable to causal inference, and to

directly investigate the effectiveness of well-specified potential interventions on modifiable components of race. The main features of the conceptual model are specification of structural racism as a key determinant of racial disparities in health and identification of a sufficient-cause interaction between black physical phenotype and structural racism.

### Aim 2

Focusing on structural racism raises the question of how to accurately measure this complex social dynamic. The second aim of the dissertation is development of a measurement model for structural racism. This work offers two contributions to the literature on measurement of racism. First, existing studies have used indices or multivariate regression modeling to estimate the effect of structural racism. Both approaches fail to account for measurement error, which may downwardly bias measures of association between structural racism and health outcomes. The latter approach estimates the independent effect of each indicator of structural racism, which fails to account for the intercorrelated and mutually reinforcing effects of dynamics like educational and employment discrimination. Here, I use confirmatory factor modeling, which minimizes bias from measurement error and takes account of the intercorrelated nature of indicators of structural racism.

I elected to measure structural racism at the county level because it is the one at which many institutional processes relevant to structural racism operate. For example, school systems responsible for setting policies that may lead to school segregation operate on the county level, rather than that of neighborhood or state. I identify five

domains relevant to structural racism: employment, education, housing, criminal justice and health. The measures fall into two groups: indicators of residential and educational segregation that measure subareal departure from the black and white proportions in county population, and prevalence ratios of the form $P_{white}/P_{black}$, where $P_{white}$ is the proportion of whites in a county experiencing an event and $P_{black}$ is the proportion of blacks experiencing the event. I used counts of non-Hispanic whites and non-Hispanic blacks when available. When a count of non-Hispanic blacks was unavailable, we used the count of all blacks. Data sources included the U.S. Census American Community Survey, the Common Core of Data collected by the U.S. Department of Education, health disparity information collected by the Dartmouth Atlas, and a census of correctional institutions administered by the federal Bureau of Justice Statistics.

Confirmatory factor modeling yielded a model with one indicator for each of the five domains of structural racism. The model exhibited acceptable fit on the confirmatory fit index, the Tucker-Lewis index, the root mean square error of approximation, and the standardized root mean square residual. I specified a nomologic network of several associations designed to evaluate convergent and divergent validity; all associations were in the hypothesized direction.

### Aim 3

The final aim of the dissertation is to provide a practical illustration of how the conceptual framework and measurement model developed in earlier aims may be used in empirical research. I chose to evaluate the association between county structural racism and BMI because BMI is a significant public health problem that is responsive to social

dynamics. Over the course of three decades, the prevalence of obesity in the United States more than doubled, to 36% in 2010, and an estimated $147 billion in health care costs in the United States were attributable to obesity in 2008. BMI trajectories are found BMI trajectories were ordered by social disadvantage, with educated white men having the lowest growth trajectory, and uneducated black women the highest. Research on structural racism and obesity has to date been limited to studies addressing a single domain—housing discrimination— indicating that content validity of existing measures is an open question.

The Aim 3 analysis relies on self-reported BMI collected as part of the Behavioral Risk Factor Surveillance System (BRFSS), an annual survey administered by the federal Centers for Disease Control and Prevention in cooperation with state health departments. Self-reported race was also obtained from BRFSS. Weighted BRFSS data are designed to yield state-level estimates. Because the Aim 3 target population is not any particular state, but rather adult residents of counties included in both the modeling of county structural racism and the BRFSS sample frame, I elected not to use the BRFSS-supplied survey weights. I modeled the structural racism/BMI relationship using a linear regression with a random effect for county identity. I specified models that included adjustment for baseline covariates (age, gender, race); baseline covariates plus structural racism; baseline covariates, structural racism, and an interaction term between black race and structural racism. Finally, to evaluate the sensitivity of our results to common macro-level confounders, I specified a model with an additional three covariates (county rurality, county median income and Census region).

I found CSR was associated with decreased BMI in the general population, but that specifying an interaction term for race, sex and CSR significantly changed the dynamics of the model. Predicted BMI for white males was substantively unchanged across levels of CSR, while BMI for white women fell 0.4 $kg/m^{2\ with}$ a change in CSR from -1SD to 1 SD. For black males, a similar change in CSR resulted in a BMI increase of 0.4 $kg/m^2$. Black females reported the highest BMI across all levels of CSR, with BMI rising 0.1 $kg/m^2$ as CSR increased from -1 SD to 1 SD. The interaction term was statistically significant ($p<0.05$). Coefficients for level 2 covariates other than CSR, and for age, were substantively the same as in the other models.

Together, the three papers open some new possibilities for research on race and health. By defining a conceptual framework that clarifies the components of race, and situations them within the potential outcomes framework, the work paves the way for additional inquiries into the causal effect of the components of race. By developing a measurement model for structural racism, I provide a validated exposure scale that can be used to deepen our knowledge of the effect of structural racism on health and other outcomes. The framework and the measurement model together underscore the importance and feasibility of designing and testing interventions on structural racism. Finally, the work with structural racism and BMI provides an illustration of how the framework and measurement model might be used in tandem to research health disparities, and raises the possibility that intervening on CSR could eliminate the racial disparity in BMI for men.

# Chapter 1

# Rethinking race and health: A new conceptual framework

## Abstract

Compared to whites in the United States, blacks experience heightened risk of many diseases as well as worse outcomes after medical treatment. These disparities are a major focus of epidemiologic investigation and public health interventions. Efforts to understand and ameliorate disparities have been complicated by two related issues: the absence of a precise definition of the biologic and social dynamics represented by the race variable, and the purportedly immutable nature of race, which hinders application of the potential outcomes framework to questions of race and health. We propose a conceptual framework separating race into distinct components, including area-level structural racism, interpersonal racism, cultural adaptation to racism and racialized group membership. We identify access to resources and stress as two important mechanisms by which structural racism affects health. Racialized group membership (black or white) and structural racism interact to produce differing effects on access to resources and psychosocial stress across groups. We conclude by exploring how the framework might facilitate further research into the health effects of structural racism, and address methodologic challenges to the framework.

## Introduction

Compared to whites in the United States, blacks experience heightened risk of many diseases (Kochanek, Anderson, and Arias 2015; DeSantis et al. 2016; Havranek et al. 2015), as well as worse outcomes after medical treatment (Lucas et al. 2006; Stone et al. 2013). Those differences, collectively known as health disparities, are a major focus of epidemiologic investigation and public health interventions.

Efforts to understand and ameliorate disparities have been complicated by two related issues that are the focus of this paper. When investigators evaluate differences in disease incidence and treatment outcomes by race, it is unclear what the "race effect" actually signifies. Research interest in race has, over the past several decades, focused variously on race as an indicator of a genotype associated with poor health, a marker of exposure to cultural and behavioral patterns that cause ill health, and a proxy for exposure to structural and interpersonal racism. The lack of a precise and widely accepted accounting of the social and biologic dynamics represented by race is a key obstacle to understanding the role of race in etiology, and to the pursuit of effective interventions.

Inference regarding the causal effect of race is another vexing area for researchers. The potential outcomes framework (Rubin 1974), the primary means by which etiologic inference is made in modern epidemiology, specifies that a causal interpretation to study findings may only emerge from a design featuring a plausible intervention. Some have argued that this implies that one cannot estimate the causal effect of race on a health outcome, given that race is typically deemed an immutable trait (Holland 1986; Kaufman and Cooper 1999).

For this reason, we suggest that a conceptual model separating race into distinct components would allow researchers to more clearly assess which components are modifiable and thus amenable to analysis under the potential outcomes framework, and to investigate the effectiveness of potential interventions. This decomposition is similar in some ways to the approach offered by VanderWeele and Robinson (2014), but differs in the specific components identified, their interrelations, and the implications for causal inference. While we term this a conceptual framework, we will address issues of causal inference, and provide a directed acyclic graph (Greenland, Pearl, and Robins 1999) covering portions of the framework.

The proposed framework (Figure 1) positions area-level structural racism as the key determinant of health. Structural racism affects health through stress and differential access to resources. The effect of structural racism on differential access to resources is mediated by cultural traits. Racialized group membership (black or white) and structural racism form an interaction in which the effect of structural racism on health varies across groups. The discussion below focuses on clearly defining each of these components, as well as assessing modifiability and potential relationships with other components in the model. We then discuss implications of, and research questions suggested by, the framework.

The framework focuses on a more precise specification of the dynamics underlying black/white health disparities, because blacks in the U.S. have experienced more frequent and severe structural and interpersonal racism than other minorities in the United States (Williams and Williams-Morris 2000). We recognize that health disparities

involving other groups represent pressing public health concerns, and note that the framework, with modifications, could be useful in research involving structural racism directed against groups other than blacks.

## Components of the framework

The framework is organized around the importance of racism to health and health disparities. Racism is increasingly recognized as a fundamental cause of health inequities (Phelan and Link 2015; Gee and Ford 2011). The fundamental cause theory identifies potential determinants of health that are linked to multiple health outcomes through the action of numerous mechanisms that change over time. It explains why race has been an important determinant of health in the United States since the nation's founding, despite large changes in the diseases responsible for mortality and morbidity, and the interventions available to fight them. In short, the theory holds that whites in the U.S. have always had preferential access to knowledge, medical care, and financial resources to secure better health outcomes. One example of this dynamic, regarding difference in trajectory of AIDS mortality by race, was documented by Rubin et al. (Phelan and Link 2015). Absent the fundamental cause theory, one might have expected the advent of effective preventative treatments for AIDS to eliminate the large racial disparity in AIDS mortality. However, the mortality gap increased steeply after the development of highly active antiretrovirals, because they were expensive and thus more widely available to whites than to blacks. For public health purposes, a key component of the fundamental cause theory is that interventions on mediators that lie between fundamental causes and health outcomes will be ineffective in reducing disparities. Thus, a conceptual framework

that clearly specifies the role of racism, and facilitates intervention upon it, is paramount.

We address two types of racism: structural and interpersonal. Gee and Ford (2011) define structural racism as "the macro level systems, social forces, institutions, ideologies, and processes that interact with one another to generate and reinforce inequities among racial and ethnic groups." Interpersonal racism is recognized as prejudice and discrimination directed toward members of a disfavored racial group by individuals in the dominant group (Jones 2000).

Our framework specifies interpersonal racism as a downstream effect of structural racism, because structural racism often reinforces interpersonal racism through the provision of differential access to resources based on race. For example, if some whites perceive blacks as criminals, this may foster interpersonal racism. Structural racism is an upstream determinant in this dynamic because it results in criminal justice policies that preferentially convict and imprison blacks, thereby bolstering the perception of blacks as criminals. Gee and Ford argue that associations between race and health outcomes would likely remain even if interpersonal discrimination were eliminated because structural racism could still persist. The framework positions psychological stress as a health determinant downstream of both structural and interpersonal racism, as both forms of racism have been shown to create stress reactions detrimental to health. (Bailey et al. 2017) Importantly, while there are practical challenges involved in modifying structural and interpersonal racism, there is no theoretical reason as to why they cannot be modified. Thus, both are suitable for consideration under the potential outcomes framework.

## Access to resources

Structural racism results in health disparities primarily by constraining access to resources for blacks. Because fundamental cause theory specifies that the particular resources involved, and the mechanisms by which they act on health, change over time (Phelan and Link 2015), the framework does not list particular resources. However, it may be informative to discuss the mechanisms by which structural racism is currently thought to affect health. Blank et al. (2004) specify five domains in which racial discrimination operates: employment, housing, health, criminal justice and education. Bailey at al. (2017) offer a similar collection of pathways through which structural racism operates: housing, education, employment, earnings, benefits, credit, media, health care, and criminal justice. These pathways are non-exclusive and interlocking. For example, new parents who are black may be unable to obtain adequate housing in a safe neighborhood due to housing discrimination. This may affect their child's health in myriad ways, including potential exposure to lead, diminished educational achievement due to differential resources available in a segregated school system, the physical and psychological effects of exposure to violence, and exposure to a neighborhood with limited employment opportunities and expectations. The latter may further diminish educational opportunities and create risk for involvement in the criminal justice system. All of these dynamics may adversely affect health of the child, as can diminished access to health resources due to employment and housing discrimination.

## The role of culture

Culture is viewed by some as an upstream and important determinant of health outcomes. And while culture does shape the behavior of individuals, it is important to

note that culture, in turn, may be shaped by structural racism. Wilson (2009) offers a definition of cultural traits that is useful in the current context: “shared outlooks, modes of behavior, traditions, belief systems, worldviews, values, skills, preferences, styles of self-presentation, etiquette, and linguistic patterns—that emerge from patterns of intra-group interaction in settings created by discrimination and segregation and that reflect collective experiences within those settings.”

Because it is not always clear how, on a practical level, structural racism might give rise to cultural traits and individual behaviors, we offer an example drawn from the educational context. Jeynes (2007) and others have found a positive association between parental involvement and student achievement. Desimone (1999) found black parents scored lower on some measures of involvement, and noted that diminished involvement has previously been viewed as a cultural trait specific to black parents. Connecting these dots without considering the determinants of cultural traits might lead one to conclude that diminished parental involvement is an independent cause of lower educational achievement, and subsequently poor health, among blacks. If, however, one closely examines the processes by which cultural traits develop, a different conclusion presents itself. Neckerman (2010) investigated the development of segregated city schools over the course of the 1900s in Chicago, finding that what might initially appear as cultural traits were instead sequelae of institutional policies: “The district's history of segregation and inequality undermined school legitimacy in the eyes of its black students. As a result, inner-city teachers struggled to gain cooperation from children and parents who had little reason to trust the school.” In keeping with these observations, the framework recognizes

that cultural traits are in part caused by, and mediate the effect of, structural racism on access to resources and health status. While it is difficult to modify culture, again there is no theoretical impediment to interventions on culture, and thus this component is suitable for consideration under the potential outcomes framework.

## The importance of racialized group membership

Any framework addressing race and health must carefully consider the role of race. Here, we focus the discussion on racialized group membership to draw attention to the socially constructed nature of race. Bonilla-Silva (1997) describes the creation of racialized groups as the assignment of social position to people sharing a set of arbitrarily identified physical characteristics: "Actors in racial positions are there not because they are of X or Y race but because X or Y has been socially defined as a race. The phenotypical characteristics of the actors are usually … used to denote racial distinctions." The social definition of race is used by individuals in combination with physical characteristics to place themselves into a racialized group, and by others to identify members of a racialized group. Accordingly, we include physical phenotype and black racialized group membership in our framework, and specify that structural racism and phenotype interact to determine the racialized group membership of individuals.

In addition to this dynamic, a second process links racialized group membership and structural racism. When members of a racialized group interact with social institutions and policies, structural racism influences the results of those interactions through differential treatment. It is important to note that this differential treatment need not involve explicitly racist policies and does not require identification of the racialized

group membership of particular individuals to operate. For example, the exodus of large employers from urban areas that has occurred over the past few decades has resulted in diminished employment for blacks, who disproportionately live in the urban areas that have been abandoned. (Wilson 2009) We know of the racialized nature of this policy only because we retrospectively observe that those who have identified themselves as black on a Census form are more likely to live in the recently vacated urban areas.

These dynamics may also operate in more explicitly racialized ways that do not require direct observation of racialized group membership. For example, in 2012 Wells Fargo bank paid $175 million to settle a lawsuit filed by the federal Department of Justice that accused the bank of a pattern of discrimination that involved channeling minority loan applicants into subprime loans, even when they were qualified for better loans (Rothacker and Ingram 2012). Wells Fargo was aware of the self-reported race of each applicant on the mortgage paperwork, but underwriters did not have personal contact with the applicants and thus had no opportunity to observe their physical characteristics (United States vs. Wells Fargo 2012).

When considering group membership within the conceptual framework, several options arise. One would be to ignore it and simply estimate the total population health effect of exposure to higher levels of structural racism. This is problematic because it does not allow for an understanding of how structural racism shapes health disparities, and does not provide visibility into how interventions might affect disparities. One might also consider simply restricting studies to participants identified by themselves or others as black, and estimating the estimate excess risk for those exposed to higher levels of

structural racism. This approach is easy to understand and implement. But it reduces statistical power, and again results in diminished information about the relationship between structural racism and health. Is structural racism protective for whites and harmful to blacks, or harmful to all members of a population? In order to fully understand these dynamics, one must examine structural racism and racialized group membership together.

### Is racialized group membership modifiable?

Prior to situating group membership in the framework, we consider the degree to which modification is plausible. As discussed above, race is typically construed as a non-modifiable characteristic. However, one benefit of the current framework is that it allows us to separately consider modifiability of each of the components of race. There are a number of ways in which group membership is modifiable, with the specifics depending on the research context. In situations similar to the Wells Fargo example described above, group membership may be directly modifiable. One could, for example, provide study subjects with a randomized value for racialized group membership and direct them to file a mortgage application as they normally would, substituting the randomized value for their true race.

In situations where structural racism is expected to operate based on visible identification of group membership, other strategies become relevant. One was outlined in *Black Like Me* (1961), in which the white author John Griffin underwent pharmacologic and ultraviolet light treatment to darken his skin, and wrote about the social and economic effect of "setting" his group membership to black. This type of

intervention blocks the effect of physical phenotype on one's racialized group membership. Such interventions would lend a causal interpretation to the prospective effect of setting a subject's group membership from white to black, or vice versa.

These types of intervention might be suitable, if ethical considerations could be managed, for a small study. But turning white people black, or vice versa, is not an effective public health intervention, whether accomplished on paper or in the flesh. Because the physical experiments considered here are interventions targeting individuals, it is likely to be more expensive and challenging to maintain than those focusing on larger social, environmental or policy features (Rose 1985). Eliminating natural variation in skin color to blunt the effect of a social system that causes ill health for people of color raises ethical concerns that are surely unresolvable. However, the potential outcomes framework does not require that an intervention be achievable on the population level to yield a causal estimate, so we proceed with the understanding that setting the racialized group membership may sometimes be feasible, but will often not be the main focus of researchers.

## Racialized group membership as a direct cause of health outcomes

Including group membership in a causal model demands further consideration of the relationships linking it to other variables of interest. One hypothesis would be that group membership is a direct cause of health outcomes. However, as discussed earlier, genotypic (Nei and Roychoudhury 1974) and epidemiologic (Cooper et al. 1997) investigations provide empirical justification for the understanding that racialized group membership is not a meaningful biological categorization, based as it is on assignment of

social value to arbitrarily selected physical features. (Bailey et al. 2017) Thus, we have no basis for hypothesizing that group membership itself is an independent cause of disease. In rare cases, the physical phenotype underlying racialized group membership can itself be a cause of disease. One example is skin cancer, where absence of melanin in lighter skin is associated with higher incidence. (Elder 1995) However, it is likely that physical traits giving rise to racialized group membership are directly associated with a limited group of diseases, mostly in the area of dermatology. Thus, we do not further consider this issue.

### Group membership in sufficient-cause interactions

Finally, we consider the possibility that racialized group membership interacts with structural racism to cause poor health for blacks. This arrangement is, in one sense, obvious to the point that it borders on tautology. A social system that is defined by differential treatment of blacks will surely produce different outcomes for blacks. However, modeling this dynamic in a way that relates to the potential outcomes framework and facilitates causal inference is work that has not yet been accomplished.

VanderWeele (2009) specifies several types of interaction under the potential outcomes framework. We consider the most relevant of these: causal interaction, sufficient cause interaction, and effect heterogeneity. Any two causes of a single variable will, by definition, interact on that variable, in that the association of one cause with the outcome changes across levels of the other cause. Specifying this type of causal interaction between racialized group membership and structural racism would, however, require a rationale for how racialized group membership represents an independent cause

of health. As previously discussed, this is generally not the case.

A more relevant possibility would be that the group membership and structural racism are synergistic causes (Greenland and Poole 1988). In this scenario, social categorization as black does not cause poor health unless a black subject is also exposed to an elevated level of structural racism. The logic of this arrangement is compelling: There is no mechanism by which structural racism could affect health if there were no blacks for the social system to operate upon. By the same token, without exposure to structural racism, black racial group membership has no relevance to health.

In the additive case, on which we focus to simplify the discussion, interpretation of a sufficient-cause relationship under the potential outcomes framework involves calculating the excess risk of poor health with both black racial group membership and exposure to structural racism less the risk incurred with the exposures separately. In practice, obtaining a causal estimate for this set of exposures requires that measurement of each be free of confounding (VanderWeele 2015). As Figure 1.2 illustrates, this will be difficult to accomplish in some instances. The figure provides a DAG illustrating many of the dynamics represented in the full conceptual model, along with additional nodes reflecting relationships between racialized group membership and structural racism over two generations. Of note are two edges, one reflecting the genetic relationship between parental group and subject group, and one reflecting the association between parental exposure to structural racism and subject exposure, due to the frequent co-location of parents and offspring. To estimate the causal effect of a sufficient-cause interaction between structural racism and racialized group, one would need to control for

parental group and parental exposure to structural racism.

### Effect heterogeneity

In some studies, this control will be feasible. If it is not, the association between racialized group and health may be interpreted as one of effect heterogeneity (VanderWeele 2009). In this event, the investigator acknowledges that interventions on racialized group membership are not in view, and the association between group and health takes on a descriptive cast. Because it is unclear whether change in the effect of structural racism across categories of racialized group is due to the group membership itself, or to a common cause shared by group membership and the outcome, the effect on the outcome of changing the group membership remains unknown. However, as we have noted, the primary public health interest lies not in changing the racialized group membership of individual subjects, but rather in understanding how interventions on structural racism are likely to eliminate health disparities. Because of this, analyses interpreted as effect heterogeneity are likely to provide useful information despite their limitations.

So far, we have limited the discussion of interaction to the relationship between the two variables central to understanding race and health—structural racism and racialized group membership. However, similar interactions occur with components in the framework downstream of structural racism. Racialized group membership modifies the effect of contact with a person who subscribes to racist ideology; blacks are likely to experience this interaction as more threatening than whites. Similarly, the development of black cultural traits in response to structural racism may not lead a white person to adopt

those traits, while the effect is different for blacks.

## Discussion

The goal of the current paper is to connect two areas of literature: one calling for more research on, and better theory and frameworks for, structural racism and health (Hicken et al. 2018; Bailey et al. 2017), and that encompassing recent developments in the area of causal inference and interaction (VanderWeele and Robinson 2014; VanderWeele 2009, 2015). Our treatment of race in the context of health disparities separates race into several components: Structural racism, interpersonal racism, cultural adaptation to structural racism, access to resources, individual behavior, psychosocial stress, and physical characteristics. We argue that racialized group membership interacts with structural racism, interpersonal racism, and cultural adaptation to alter the way in which they affect health. We note that taken separately, the components of race are generally modifiable in a prospective sense and thus suitable for analysis under the potential outcomes framework. Because structural racism is a fundamental cause, intervention on mechanisms lying between structural racism and health will not be as effective as intervening on structural racism itself.

While modifiability remains an important criteria for many investigators, others (Pearl 2010; Glymour and Spiegelman 2017) have argued that the no manipulation, no causation dogma is misguided. Pearl (2010) advocated for a focus on whether descendants of an exposure receive signals from the exposure, not whether the exposure is modifiable. The intervention required to achieve a causal contrast is on signal receptivity rather than on the value of the exposure. Thus, in the context of the current

paper, the question would not be whether race is modifiable, but rather whether the social system downstream from race reacts to information about race. Our research facilitates the type of causal contrast Pearl proposed by specifying that racialized group membership represents the joint action of individual physical phenotype and a system that is designed to assign social value to (or receive signals about) phenotype.

The framework has a number of other strengths. We propose specific relationships between the components of race and health outcomes that may lead to causal inference using the potential outcomes framework. In many cases, these relationships can be empirically tested. Such testing may lead to useful new lines of research. For example, the nature of the relationship between structural racism, interpersonal racism and health is not well covered in the current literature.

One important implication of our work is that structural racism must be accurately measured to understand the dynamics underlying race and health. Existing work on structural racism has focused on single-indicator measures and indices, which are likely to result in measurement error and poor content validity. The question of geographic scale becomes relevant here, because it is not immediately clear whether structural racism affects health most significantly on the neighborhood, county, state or national level. Another possibility is that structural racism operates on all of those levels, and that multilevel measurement approaches are necessary to adequately account for those dynamics.

Designating appropriate control units when gauging the effect of structural racism will likely be important. Because the United States has a pervasive history of racial

subjugation, it is unlikely that unexposed persons or areas can be found within the country. The problem is illustrated in an example given by Rose (1985), who discussed the relationship between soft drinking water and heart disease. If one evaluates cardiovascular deaths in Scotland, there is no association with drinking water. This is because all subjects receive soft tap water and there is no variation in the exposure. If one expands the analysis to other regions receiving hard tap water, the association between soft water and cardiovascular risk becomes clear. Evaluating the effect of structural racism may require cross-national analysis, which raises additional methodologic concerns.

Another implication of the framework relates to the advisability of adjusting for socioeconomic status and other variables, such as housing, neighborhood quality, and educational attainment, indicative of access to resources. A unique feature of our conceptual model is that it specifies a mediating role for these variables. These variables are frequently considered to be confounders of any relationship between race and health outcomes. However, because of the powerful effect that structural racism has on socioeconomic status and education, we suggest that, in the context of studies evaluating the effect of structural racism on health, adjusting for these variables will result in biased measurement of the effect of structural racism.

Finally, we consider the requirement for controlling for parental exposure to structural racism when estimating the causal effect of structural racism. Data on location of parental residence are rarely collected in administrative and cohort data, so this type of adjustment will often be infeasible. Importantly, this is a limitation affecting many

epidemiologic studies. We could easily supply a causal diagram for lung cancer, for example, that would position parental smoking as a common cause of descendant smoking and childhood exposure to smoke, both of which might cause cancer. This possibility was not an impediment to Doll and Hill (1950), and it should not cause others to avoid studies on structural racism, either. When adjustment is not possible, the structural racism effect will reflect the subject's exposure to structural racism as well as parental exposure to structural racism. This is an obstacle to estimating the immediate effect of intervening on structural racism. However, such an analysis has scientific value on its own, in that it measures the effect of historical structural racism.

We share some of the concerns expressed by Schwartz et al. (2016), namely that the potential outcomes framework serves a socially conservative function by limiting consideration of causes to those dynamics that might be addressed through small-scale interventions operationalized as randomized controlled trials. However, we offer the proposed framework as an example of how detailed consideration of the social dynamics underlying causes seemingly outside the reach of the potential outcomes framework can, in fact, yield workable counterfactuals. While programs addressing area structural racism are surely challenging to conceive of and execute, it is possible to envision well-defined interventions. For example, one might institute a policy or law requiring county governments to measure and ameliorate racial differences across the domains in which structural racism commonly affects disparities. Bailey et al. (2017) discuss others, including Purpose Built Communities, an effort to redevelop a large number of neighborhoods of concentrated disadvantage with mixed-income residences and co-

located services to promote health, education and employment. We acknowledge this type of intervention would, as argued by Schwartz et al., estimate the effect of a particular intervention rather than the effect of the previously established structural racism. This is a limitation in some respects, a benefit in others. We expect that achieving clarity about the potential effect of interventions on structural racism will ultimately result in more and better interventions.

**Figure 1.1: Conceptual framework—structural racism, racialized group and health**

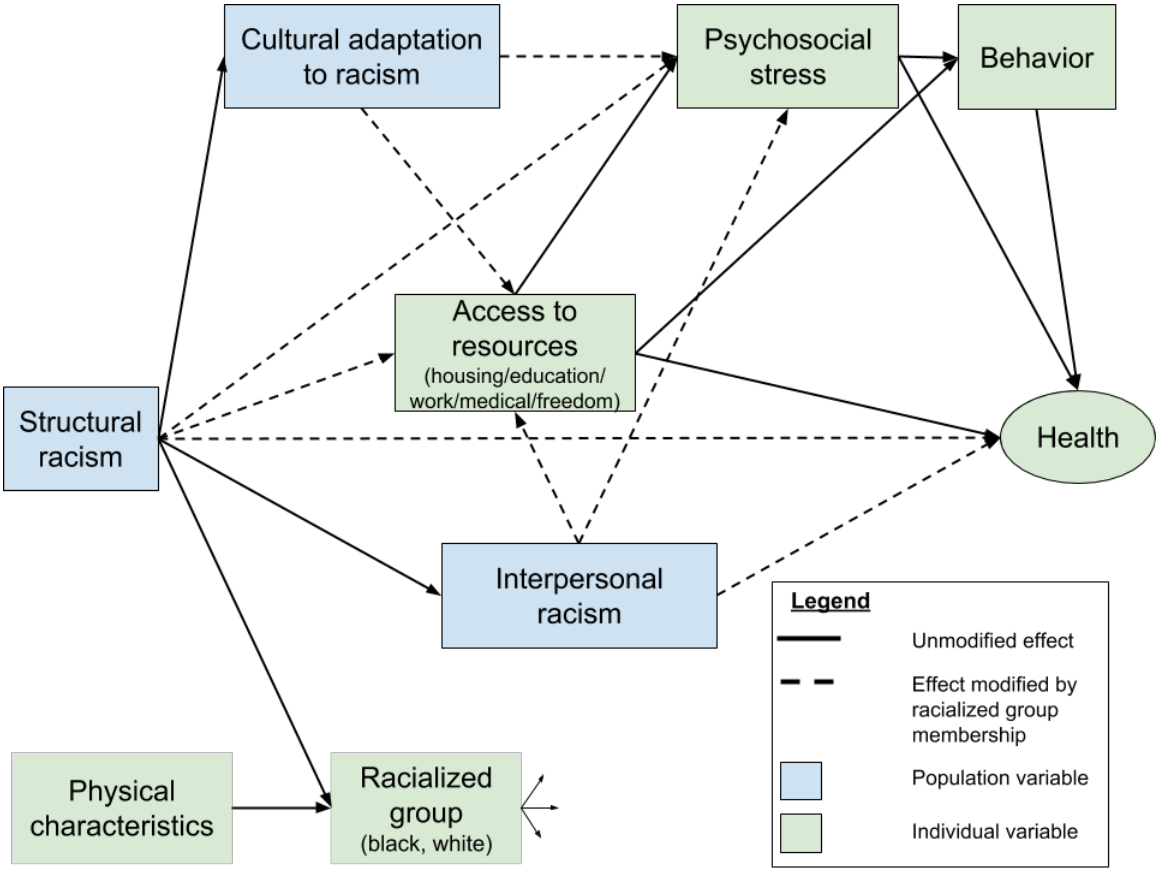
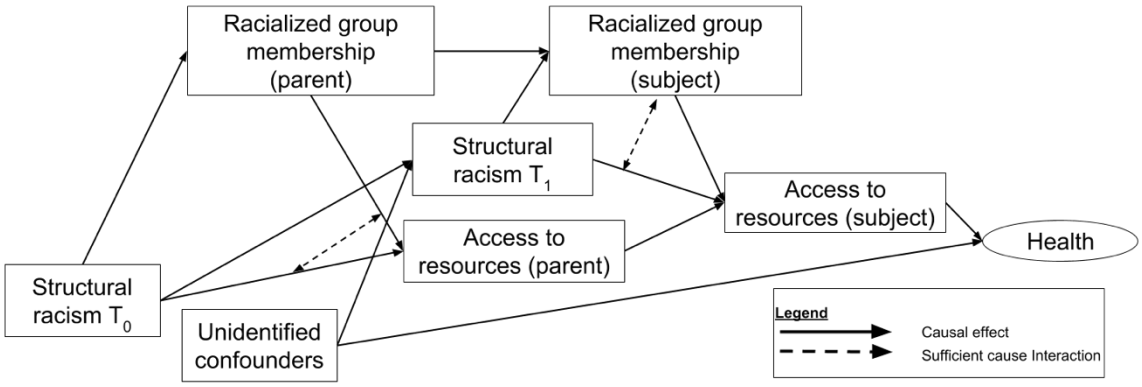


**Figure 1.2: Directed acyclic graph, bi-generational structural racism, racialized group membership and health**




**The graph shows causal relationships involving structural racism over two generations. $T_0$ represents measurement at a time prior to the birth of the subject when structural racism would have affected parents' access to material resources. $T_1$ represents measurement during subject's adulthood.**

## References


Bailey, Zinzi D., Nancy Krieger, Madina Agénor, Jasmine Graves, Natalia Linos, and Mary T. Bassett. 2017. "Structural Racism and Health Inequities in the USA: Evidence and Interventions." *The Lancet* 389 (10077): 1453–63.

Blank, Rebecca M., Marilyn Dabady, and Constance F. Citro. 2004. *Measuring Racial Discrimination*. Washington, DC: National Academies Press.

Bonilla-Silva, Eduardo. 1997. "Rethinking Racism: Toward a Structural Interpretation." *American Sociological Review* 62 (3). [American Sociological Association, Sage Publications, Inc.]: 465–80.

Cooper, R., C. Rotimi, S. Ataman, D. McGee, B. Osotimehin, S. Kadiri, W. Muna, et al. 1997. "The Prevalence of Hypertension in Seven Populations of West African Origin." *American Journal of Public Health* 87 (2): 160–68.

DeSantis, Carol E., Rebecca L. Siegel, Ann Goding Sauer, Kimberly D. Miller, Stacey A. Fedewa, Kassandra I. Alcaraz, and Ahmedin Jemal. 2016. "Cancer Statistics for African Americans, 2016: Progress and Opportunities in Reducing Racial Disparities." *CA: A Cancer Journal for Clinicians* 66 (4): 290–308.

Desimone, Laura. 1999. "Linking Parent Involvement With Student Achievement: Do Race and Income Matter?" *The Journal of Educational Research* 93 (1). Routledge: 11–30.

Doll, R., and A. B. Hill. 1950. "Smoking and Carcinoma of the Lung; Preliminary Report." *British Medical Journal* 2 (4682). ncbi.nlm.nih.gov: 739–48.

Elder, D. E. 1995. "Skin Cancer. Melanoma and Other Specific Nonmelanoma Skin Cancers." *Cancer* 75 (1 Suppl). Wiley Online Library: 245–56.

Gee, Gilbert C., and Chandra L. Ford. 2011. "STRUCTURAL RACISM AND HEALTH INEQUITIES: Old Issues, New Directions." *Du Bois Review: Social Science Research on Race* 8 (1): 115–32.

Glymour, M. Maria, and Donna Spiegelman. 2017. "Evaluating Public Health Interventions: 5. Causal Inference in Public Health Research-Do Sex, Race, and Biological Factors Cause Health Outcomes?" *American Journal of Public Health* 107 (1). ajph.aphapublications.org: 81–85.

Greenland, S., J. Pearl, and J. M. Robins. 1999. "Causal Diagrams for Epidemiologic Research." *Epidemiology* 10 (1). JSTOR: 37–48.

Greenland, S., and C. Poole. 1988. "Invariants and Noninvariants in the Concept of Interdependent Effects." *Scandinavian Journal of Work, Environment & Health* 14 (2). JSTOR: 125–29.

Griffin, John Howard. 1961. *Black like Me*. Houghton Mifflin.

Havranek, Edward P., Mahasin S. Mujahid, Donald A. Barr, Irene V. Blair, Meryl S. Cohen, Salvador Cruz-Flores, George Davey-Smith, et al. 2015. "Social Determinants of Risk and Outcomes for Cardiovascular Disease: A Scientific Statement From the American Heart Association." *Circulation* 132 (9): 873–98.

Hicken, Margaret T., Nicole Kravitz-Wirtz, Myles Durkee, and James S. Jackson. 2018. "Racial Inequalities in Health: Framing Future Research." *Social Science & Medicine*, January. Elsevier. https://doi.org/10.1016/j.socscimed.2017.12.027.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American*

*Statistical Association* 81 (396): 945–60.
Jeynes, William H. 2007. "The Relationship Between Parental Involvement and Urban Secondary School Student Academic Achievement: A Meta-Analysis." *Urban Education* 42 (1). SAGE Publications Inc: 82–110.
Jones, C. P. 2000. "Levels of Racism: A Theoretic Framework and a Gardener's Tale." *American Journal of Public Health* 90 (8): 1212–15.
Kaufman, J. S., and R. S. Cooper. 1999. "Seeking Causal Explanations in Social Epidemiology." *American Journal of Epidemiology* 150 (2): 113–20.
Kochanek, Kenneth D., Robert N. Anderson, and Elizabeth Arias. 2015. "Leading Causes of Death Contributing to Decrease in Life Expectancy Gap Between Black and White Populations: United States, 1999-2013." *NCHS Data Brief*, no. 218 (November). researchgate.net: 1–8.
Lucas, F. L., Therese A. Stukel, Arden M. Morris, Andrea E. Siewers, and John D. Birkmeyer. 2006. "Race and Surgical Mortality in the United States." *Annals of Surgery* 243 (2): 281–86.
Neckerman, Kathryn M. 2010. "Schools Betrayed." University of Chicago Press. pu3430623_3430810. June 2010. http://www.press.uchicago.edu/ucp/books/book/chicago/S/bo5298931.html.
Nei, M., and A. K. Roychoudhury. 1974. "Genic Variation within and between the Three Major Races of Man, Caucasoids, Negroids, and Mongoloids." *American Journal of Human Genetics* 26 (4). ncbi.nlm.nih.gov: 421–43.
Pearl, Judea. 2010. "3. The Foundations of Causal Inference." *Sociological Methodology* 40 (1). Wiley Online Library: 75–149.
Phelan, Jo C., and Bruce G. Link. 2015. "Is Racism a Fundamental Cause of Inequalities in Health?" *Annual Review of Sociology* 41 (1). Annual Reviews: 311–30.
Rose, G. 1985. "Sick Individuals and Sick Populations." *International Journal of Epidemiology* 14 (1): 32–38.
Rothacker, Rick, and David Ingram. 2012. "Wells Fargo to Pay $175 Million in Race Discrimination Probe." *Reuters*, July 12, 2012. https://www.reuters.com/article/us-wells-lending-settlement/wells-fargo-to-pay-125-million-in-fair-lending-case-idUSBRE86B0V220120712.
Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5). American Psychological Association: 688.
Schwartz, Sharon, Seth J. Prins, Ulka B. Campbell, and Nicolle M. Gatto. 2016. "Is the 'Well-Defined Intervention Assumption' Politically Conservative?" *Social Science & Medicine* 166 (October). ncbi.nlm.nih.gov: 254–57.
Stone, Matthew L., Damien J. Lapar, Bartholomew J. Kane, Sara K. Rasmussen, Eugene D. McGahren, and Bradley M. Rodgers. 2013. "The Effect of Race and Gender on Pediatric Surgical Outcomes within the United States." *Journal of Pediatric Surgery* 48 (8). jpedsurg.org: 1650–56.
United States vs. Wells Fargo. 2012. U.S. District Court for the District of Columbia.
VanderWeele, T. J. 2009. "On the Distinction between Interaction and Effect Modification." *Epidemiology* 20 (6): 863–71.
———. 2015. *Explanation in Causal Inference: Methods for Mediation and Interaction*.

Oxford University Press.
VanderWeele, T. J., and W. R. Robinson. 2014. “On the Causal Interpretation of Race in Regressions Adjusting for Confounding and Mediating Variables.” *Epidemiology* 25 (4): 473–84.
Williams, D. R., and R. Williams-Morris. 2000. “Racism and Mental Health: The African American Experience.” *Ethnicity & Health* 5 (3-4): 243–68.
Wilson, William Julius. 2009. “Toward a Framework for Understanding Forces That Contribute to or Reinforce Racial Inequality.” *Race and Social Problems* 1 (1). Springer US: 3–11.

# Chapter 2

# Development and validation of a scale measuring county structural racism

## Abstract

Blacks in the United States experience diminished life chances compared to whites in domains ranging from employment to health, education, and criminal justice. In recent years, researchers have focused on racism as a key determinant of such differences. This has in turn increased interest in measuring racism. Recent measurement efforts have focused on individual-level instruments measuring perceived racism, and on regression models that incorporate one or more proxies of institutional or structural racism as independent variables. Each approach involves methodologic challenges that may lead to mismeasurement of racism. Because of those issues, we elected to measure structural racism at the community level by evaluating differences in institutional treatment of blacks and whites, using a factor analysis approach that accounts for measurement error and the correlated nature of potential indicators of structural racism. Our approach improves upon the content validity of previous work by measuring racism with indicators in five domains: housing, education, employment, healthcare and criminal justice. We estimated a structural racism factor score for 1,787 counties with a black population of greater than 500, representing more than half of the counties and 92% of the population in the country. The model demonstrated acceptable fit across a panel of statistics designed to identify poor-fitting factor models, and performed well in a construct validity analysis that involved construction of a detailed nomologic network. The resulting county estimates of structural racism can be used to investigate the association between racism and racial disparities across a variety of domains.

## Introduction

Blacks in the United States experience diminished life chances compared to whites in domains ranging from employment to health, education, and criminal justice (Bonilla-Silva 2009; Chen and Nomura 2015; Nunley et al. 2015; DeSantis et al. 2016; Reardon, Yun, and Eitle 2000). Prior to the development of genome-sequencing studies that decomposed genetic variation into within- and between-race components, investigators often attributed black-white differences to biological factors. In recent decades, though, epidemiologic analyses (Cooper et al. 1997), and molecular studies (Nei and Roychoudhury 1974), have proven this theory false. Another theory is that individual behavior, sometimes driven by culture, is a key cause of black-white differences. More recently, researchers have begun developing evidence that racism, rather than race or culture, is a powerful explanation for racial differences. (Dressler, Oths, and Gravlee 2005)

These developments have led to increased interest in measuring racism. Most of this effort has been focused on individual-level instruments that measure racism as perceived by survey subjects. A recent systematic review identified 26 such measurement instruments. (Bastos et al. 2010) This focus on perceived racism raises a number of methodologic issues. (Priest et al. 2013) It may be unclear, from the viewpoint of a minority group member, whether unfavorable treatment is attributable to racism or to a more benign explanation, so not all interactions driven by racism may be perceived as such. (Major, Quinton, and McCoy 2002) Subjects who are members of multiple low-status groups, such as black men who have sex with men, may be unable to determine whether hostile encounters are attributable to racism or other types of bias. (Bastos et al.

2010) Exposure to racism may be mitigated by coping strategies that alter a respondent's perception of events or interest in reporting them.

Finally, there is the question of measurement level. Structural racism is defined as the interconnected societal and institutional practices that result in differential treatment by race. Following Gee and Ford (2011) and Jones (2000), we argue that measuring and intervening on structural racism is more important than similar efforts directed at interpersonal racism. Structural racism represents an upstream determinant of perceived racism and racial disparities, and may be a "fundamental cause" (Link and Phelan 1995; Phelan and Link 2015) of racial disparities in health and other domains. Because of this, it is possible that negative effects of race would continue to exist even if interpersonal racism were eliminated. Further, eliminating structural racism may result in reductions in interpersonal racism, as it reduces the state-sanctioned nature of differential treatment by race, and over time eliminates differences in health, education and socioeconomic status between racial groups.

While measuring structural racism is paramount to understanding racial disparities, this goal is not well served by currently available instruments of perceived racism. Some of these instruments contain items or subdomains related to structural racism (Green 1995; Utsey 1999), while others do not explicitly measure structural racism. (Taylor, Kamarck, and Shiffman 2004) Regardless, structural racism is a societal, rather than individual, phenomena, and thus is not adequately measured at the individual level. A related, albeit more practical concern, is that individuals exposed to structural racism may not be aware of it. A black person has no reason to suspect their mortgage

application was denied because of race, absent knowledge of how the bank handled applications filed by similarly situated whites.

Because of the issues noted above, researchers have started measuring structural racism at analytic levels above the individual respondent. The prevalent approach to this type of measurement (Lukachko, Hatzenbuehler, and Keyes 2014; Wallace et al. 2017; Zhou, Bemanian, and Beyer 2017; Mendez, Hogan, and Culhane 2011, 2014; Jacoby et al. 2017) involves investigating structural racism by assembling a panel of indicators, generally one per domain, and regressing the outcome of interest against the panel of indicators. Modeling multiple domains of a complex latent phenomena as independent predictors has several drawbacks. Due to concerns about multicollinearity, this approach makes it difficult to include multiple indicators in a particular domain, so the analysis loses the accuracy provided by a full suite of indicators. The estimates resulting from this modeling approach are of the independent effect of a particular variable, holding all other variables constant. However, to understand the impact of structural racism on disparities, one needs an estimate of the effect of all of the domains operating in concert on the outcome, while accounting for the intercorrelation between domains. This requirement points one in the direction of factor modeling, the benefits of which include estimates that account for the intercorrelated nature of indicators, unbiased estimates in the presence of measurement error, a robust set of fit statistics specific to the goal of measuring latent constructs and the ability to specify an error structure for the indicators. Finally, studies to date have not specified a set of indicators with strong content validity across all domains of structural racism.

To accurately gauge the effect of structural racism on key social outcomes, to evaluate potential interventions, and to understand the relationship between interpersonal and structural racism, one must measure structural racism accurately. To our knowledge, a validated scale measuring multiple domains of structural racism has not been developed.

Accordingly, the central thrust of this paper is to develop and validate a scale measuring structural racism at the county level in the United States. We seek to more fully understand the relationship between the latent construct of structural racism and the observed indicators caused by it, and to extract factor scores reflecting the relative level of structural racism across U.S. counties, with the goal of facilitating further research on structural racism and its association with health and other individual outcomes.

Because structural racism is a complex, latent concept that is not measurable via a single variable, we propose a confirmatory factor analysis model, relying on a robust selection of macro-level indicators of discriminatory treatment, to obtain accurate measurement of structural racism.

## Materials and Methods

We focus on structural racism at the county level. Previously, authors have presented analyses on discrimination using a variety of geographic units, including nations (Mayda 2006), states (Krieger et al. 2013), counties (Foster and Kleit 2015), and smaller localities (Lichter, Parisi, and Taquino 2012). We focus on counties for several reasons. Some candidate indicators (comparison of segregation across schools, for example) necessitate a unit of analysis larger than Census tract or neighborhood.

However, there is little doubt that policies implemented by governmental units below the state level are important drivers of discrimination. For example, school systems are often operated by county governments, which set educational policies that can encourage, or prohibit, school segregation. Many potential indicators were readily available at the county level. Finally, creating a measure based on a sub-county unit of analysis would exclude areas of the nation in which the county represents the smallest unit of government.

A scale measuring structural racism could address differential treatment of minorities in at least two ways: evaluating differences between whites and all minorities, and evaluating differences between whites and a specific minority group. Williams and Williams-Morris (2000) note that blacks in the United States have been subjected to a level of structural and interpersonal racism that far exceeds that experienced by members of other minority groups. Housing covenants explicitly targeted blacks, rather than Hispanics and Asians. Given this background, it is difficult to conceive of an approach to measuring structural racism that simultaneously and meaningfully accounts for a given county's treatment of all relevant racial and ethnic groups. The difficulty becomes more apparent when one considers the degree to which structural racism directed against Hispanics and Asians may be complicated by anti-immigrant bias. We elected to measure white vs. black county structural racism (CSR), while noting that the methods we apply could be readily adapted to measure other forms of structural racism.

### Development of the CSR scale

Blank et al (2004) identified five domains of racial discrimination: employment,

housing, education, criminal justice and healthcare. We sought candidate indicators that measure differential black/white treatment across these five domains.

Researchers have constructed well-established single-domain measures of discrimination at the area level (Massey and Denton 1988), which are often employed in housing and education studies. Broadly, housing and education discrimination indicators evaluate the degree to which units (Census tracts or schools) within a county mirror the racial composition of the county. When possible, we evaluated two variants of these indicators: One in which clustering of blacks was compared to the rest of the population, and one in which blacks were compared to the white population. We found the latter to be better correlated with other indicators of CSR, and so used this approach in our modeling.

Area-level discrimination in employment, criminal justice and healthcare has been less closely studied. We sought readily available measures of differential treatment by race. We evaluated multiple indicators for each domain, and relied on modeling diagnostics to select the most informative indicators. These indicators, with few exceptions, are prevalence ratios: $P_{w=1}/P_{b=1}$, where $P_{w=1}$ is the proportion of whites in a county experiencing an event and $P_{b=1}$ is the proportion of blacks experiencing the event. When the data permitted calculations using counts of non-Hispanic whites and non-Hispanic blacks, we used those categories. In some instances, the count of non-Hispanic blacks was unavailable. In those cases, we used the count of all blacks.

Our measurement effort relies primarily on U.S. Census Bureau survey data and community-level data collected for administrative reasons (Table 1). We excluded any

county with a black population of less than 500. We sought to estimate CSR for 2009, which would provide an appropriate lag period to use the CSR measure as a predictor of recently measured racial disparities. This decision led us to use the 2007-2011 American Community Survey five-year data file, and to select a similar vintage for other files. Some files, such as the U.S. Department of Justice Census of Jails, were not available for 2009. In those instances, we used the available vintage closest to 2009.

We selected confirmatory factor analysis (CFA) to estimate CSR because CFA accommodates measurement error in observed indicators, empirically determines the degree to which each indicator is weighted in the composite estimate, and provides multiple indices of fit that aid selection of a final model. The fit statistics allow the investigator to ensure that the selected model has adequately reproduced the covariance structure of the indicator variables, while the loadings provide evidence of the reliability of individual indicators, as well as confirmation that indicators from all content areas are substantively represented in the final model.

We fit models using a robust maximum likelihood estimator, which produces unbiased estimates using data with missing and skewed indicators. We deemed a model to have acceptable fit if the comparative fit index (CFI) and Tucker-Lewis fit index (TLI) were greater than 0.95, root mean square error of approximation (RMSEA) was below 0.06, and standardized root mean square residual (SRMR) was below 0.08. (Hu and Bentler 1999)

### Construct validity analysis

We established three tests of convergent validity drawn from literature on the

causes and sequelae of structural racism. Reich (1971) presented empirical evidence that structural racism is associated with social arrangements designed to maintain income inequality. Accordingly, we hypothesized counties with higher levels of CSR would exhibit higher levels of income inequality between Census tracts, as measured by the Gini index.

Wilson (1978) argued that structural racism results in a higher proportion of female-headed black households. The reasons for this are complex and varied. Poor labor prospects for black men and increased racial disparity in incarceration rates are two of these reasons. Using the U.S. Census American Community Survey, we obtained the proportions of black and white households in each county. We hypothesized CSR would be positively associated with the proportion of female-headed black households. Lastly, we examined social capital, under the theory that processes caused by structural racism, such as the channeling of black people into neighborhoods of concentrated disadvantage, may diminish social capital in those neighborhoods. (Portes 1998; Sampson, Raudenbush, and Earls 1997) Martin and Newman (2014) determined that Census response rates are a valid proxy for neighborhood social capital. We estimated the white-black social capital differential by assigning each county resident the Census response rate of their Census block, and then created a frequency-weighted average for blacks and whites. We subtracted the black value from the white value to arrive at the average racial difference in social capital. We hypothesized that counties with higher CSR would exhibit a larger white-black social capital difference.

To assess discriminant validity, we examined associations of CSR with the

percentage of black residents in the county and the total county population. We deemed correlations of absolute magnitude less than 0.3 indicative of acceptable discriminant validity.

## Results

We identified 1,787 counties with a black population of greater than 500, representing more than half of the counties in the country. These counties accounted for 282,093,181 residents, or 92% of the total U.S. population (Table 2). Included counties were predominantly located in the South and Midwest regions. The proportion of black residents per county ranged from 2.8% in the West to 21.8% in the South. Mean (SD) housing dissimilarity index was 0.42 (0.15). Mean white/black high school graduation ratio was 1.05 (0.13) Mean incarceration ratio was 6.48 (9.31). Mean poverty ratio was 2.37 (2.14). Mean A1c ratio was 1.02 (0.08). Pearson correlation coefficients between the indicators ranged from 0.009 (poverty ratio and A1c control ratio) to 0.48 (Housing dissimilarity index and high school graduation ratio).

Factor analysis modeling resulted in a final model with one indicator from each of the five content domains, and acceptable fit statistics. (Table 3) The final model included the dissimilarity index (housing); black/white risk ratio of education less than 12th grade education; the black/white poverty ratio; the black/white incarceration ratio; and the white/black diabetes care ratio. Housing and education were loaded most highly in the final model (Figure 1). All loadings were statistically significant.

Results of our construct validity analysis are shown in Table 4. The factor scores generated from the final model were positively correlated with the Gini index of income

inequality, the proportion of black households headed by a single woman, and the white-black social capital difference. The factor scores were correlated at 0.32 with county total population and -024 with the county proportion of black residents.

## Discussion

We developed a measurement model for county structural racism that fits well to data, exhibits strong construct validity and covers five important domains of discrimination. From that model, we derived county factor scores for structural racism. These are, to our knowledge, the first local estimates of structural racism that rely on measurements across multiple domains and address the issue of measurement error.

The analysis highlights some issues in measuring structural racism. For example, initially, we sought to include a measure of black vs. white exposure to police-on-citizen violence. However, our work with county-level death certificates led us to conclude that these events are rare enough to prevent meaningful inter-race comparison in most counties. Looking at the discrimination domain, we had initially suspected, based on the work of Massey and Denton (1988), that we would be able to specify a well-fitting model that contained multiple indicators of residential segregation. This did not prove to be the case. However, the selected model includes the dissimilarity index, which is the most readily interpretable, and one of the most commonly used, segregation measures.

Our work benefited from pre-specification of a detailed nomologic network that we used to gauge construct validity. This work consisted of identifying constructs that, based on the literature, are likely sequelae of structural racism. In many cases, the effect of structural racism on these constructs is theorized to be mediated by the items used as

indicators in the CFA. One relationship not included in our pre-specified network is worthy of note. We specified that CSR would be positively associated with the proportion of black households headed by females, and found the data supported this hypothesis.

One limitation of our work is that the measurement model we developed is not directly applicable to measurement of structural racism outside of the white vs. black context. Unquestionably, other forms of racism, such as white vs. Hispanic, have figured prominently in American social affairs. We focused on white vs. black racism because considering multiple forms of racism in the same analysis would have been impractical. While white vs. black racism is the longest standing and most pervasive form of racism, the current analysis illustrates a method that could be used to measure other forms of racism.

Similarly, the analysis relied on single, rather than multiple, categories of race in deriving black vs. white indicators. The methods applied here could be used to evaluate structural racism involving various combinations of races. However, given that racism in the United States has often defined black race using the "one-drop" rule (Khanna 2010), our use of single-race categories is consistent with the phenomena under measurement. While counties represent one appropriate level of measurement for structural racism, there may be others. Our selection of this geographic scale could produce a method effect that would bias the association between structural racism and other variables of interest. This is an area in which further investigation is warranted.

A final limitation involves the question of time scale. The indicators included in

the CSR model certainly measure the degree to which contemporaneous policies result in differential treatment by race. The scale may also measure the effect of former policies that created racial disparities in a prior generation, which in turn resulted in current disparities due to differences in access to resources prior to birth and during early childhood. Untangling the complex web of causality over time is beyond the scope of the current project. Future research, however, might focus on evaluating time-varying measures of CSR, and understanding their relationship to the level of racial disparity over time.

As interest in the association between structural racism and racial disparities grows, so does the need for a measure of structural racism that exhibits strong content and construct validity. The CSR measurement model described here provides a useful way for investigators to incorporate the concept of structural racism into their research.

**Table 2.1: Description of selected county-level indicators**

| Indicator | Domain | Data Source | Database | Description |
|---|---|---|---|---|
| Dissimilarity index | Housing | U.S. Census Bureau | American Community Survey SF-1 (2007-11) | Proportion of blacks that would have to relocate to achieve even distribution. |
| High-school graduation ratio | Education | U.S. Census Bureau | American Community Survey SF-1 (2007-11) | Ratio of non-Hispanic white to black high school graduation rates |
| Incarceration ratio | Criminal justice | U.S. Department of Justice | Census of Jail Inmates, 2005 | Ratio of non-Hispanic white to black county jail incarceration |
| Poverty ratio | Employment | U.S. Census Bureau | American Community Survey SF-1 (2007-11) | Ratio of non-Hispanic white to black poverty proportions |
| Diabetes prevention ratio | Health | Dartmouth Atlas of Health Care | 2012 Atlas | Ratio of white non-Hispanic diabetics receiving appropriate A1C monitoring to black diabetics receiving appropriate A1C monitoring |

*Note: All indicators are oriented so that larger values indicate higher county structural racism*

**Table 2.2: Counties included in study, by Census region**

| | **Northeast** | **Midwest** | **South** | **West** |
|---|---|---|---|---|
| N | 162 | 332 | 1077 | 138 |
| Population, mean (SD) | 327577 (404658) | 165743 (352936) | 100681 (235066) | 475112 (1000982) |
| % black residents, mean (SD) | 6.5 (7.2) | 6.1 (6.8) | 21.8 (18.1) | 2.8 (2.8) |
| Thiel's H, mean (SD) | 4.3 (1.1) | 3.9 (1.1) | 3.1 (1.2) | 3.5 (0.7) |
| White/black HS graduation ratio, mean (SD) | 2.1 (1.0) | 2.2 (1.2) | 1.8 (0.8) | 2.1 (1.2) |
| White/black incarceration ratio, mean (SD) | 3.0 (1.0) | 2.6 (1.0) | 2.1 (0.8) | 2.2 (0.9) |
| White/black unemployment ratio, mean (SD) | 6.8 (2.1) | 7.2 (2.6) | 7.0 (2.1) | 6.4 (2.3) |
| White/black median income difference, mean (SD) | 2.0 (1.3) | 1.8 (1.3) | 1.9 (0.9) | 1.6 (1.5) |
| White/black health insurance ratio, mean (SD) | 2.3 (2.6) | 2.3 (2.3) | 1.6 (1.8) | 1.5 (2.4) |

**Table 2.3: Correlation matrix, confirmatory factor items**

| | **Housing dissimilarity index** | **HS graduation ratio** | **Incarceration ratio** | **Poverty ratio** | **A1C control ratio** |
|---|---|---|---|---|---|
| Housing dissimilarity index | 1 | | | | |
| HS graduation ratio | 0.475 | 1 | | | |
| Incarceration ratio | 0.279 | 0.152 | 1 | | |
| Poverty ratio | 0.149 | 0.185 | 0.119 | 1 | |
| A1C control ratio | 0.139 | 0.0899 | 0.0731 | 0.00923 | 1 |

**Figure 2.1: Measurement Model of County Structural Racism**

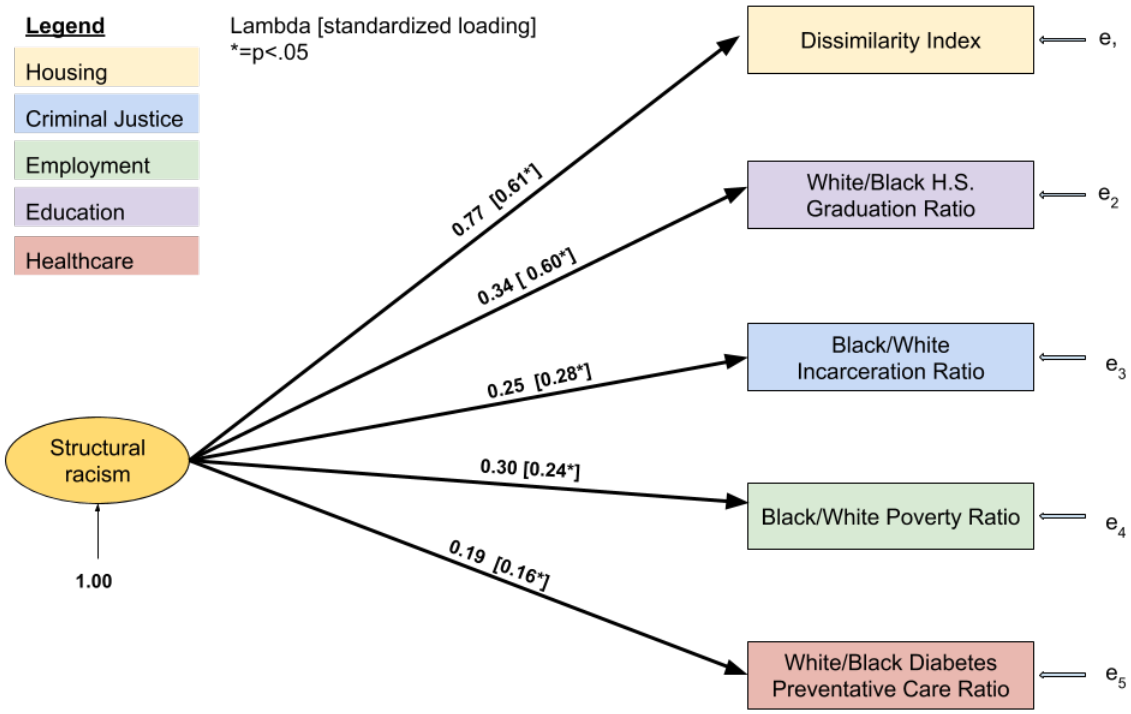


**Table 2.4: Confirmatory factor analysis fit statistics**

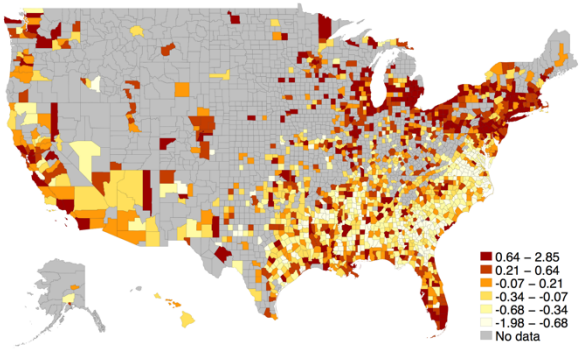
| **Fit statistic** | **Estimate** |
|---|---|
| Confirmatory fit index | 0.979 |
| Tucker-Lewis fit index | 0.958 |
| Root mean square error of approximation | 0.026 |
| Standardized root mean square residual | 0.022 |

**Table 2.5: Association of latent variable with items in nomologic network**

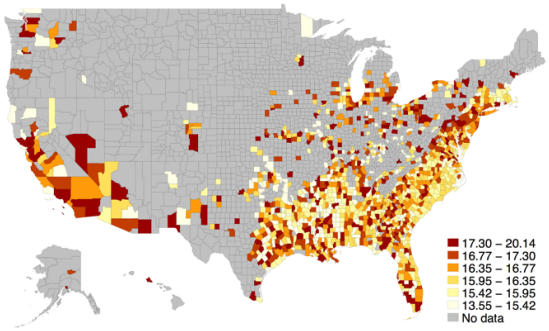
| Item | Rho |
|---|---|
| Percentage of black households headed by females | 0.126 |
| Percentage of white households headed by females | 0.059 |
| White-black difference in social capital, as measured by Census response rate | 0.388 |
| Gini coefficient, income | 0.067 |
| County population | 0.32 |
| Percent black | -0.24 |

**Figure 2.2: Choropleth maps, quintiles of factor score and items**

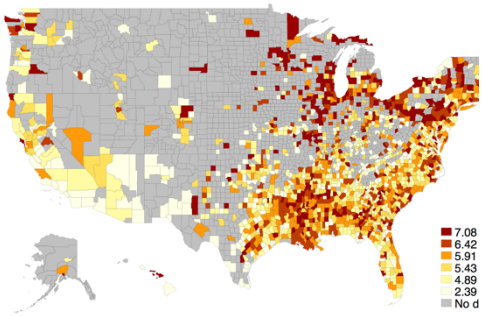
Quintiles of structural racism



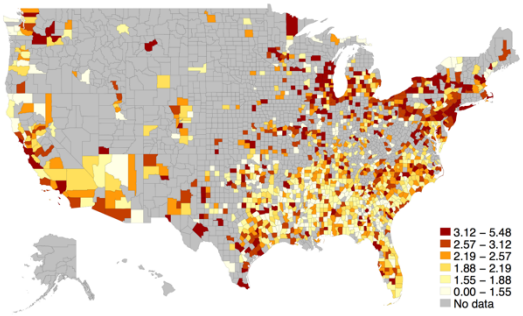

Quintiles of A1c control ratio



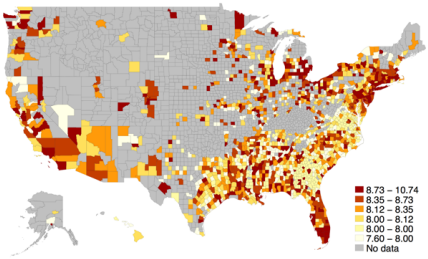

Quintiles of poverty ratio



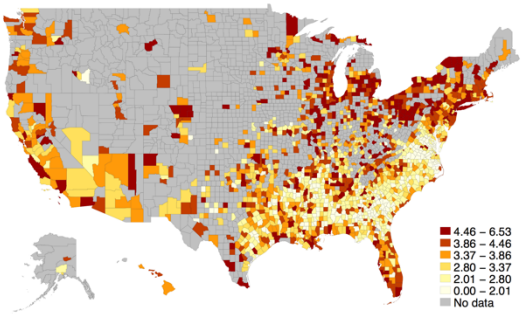

Quintiles of incarceration ratio




Quintiles of high school graduation ratio




Quintiles of housing dissimilarity index



## References


Bastos, J. L., R. K. Celeste, E. Faerstein, and A. J. Barros. 2010. "Racial Discrimination and Health: A Systematic Review of Scales with a Focus on Their Psychometric Properties." Social Science & Medicine 70 (7): 1091–99.

Blank, Rebecca M., Marilyn Dabady, and Constance F. Citro. 2004. *Measuring Racial Discrimination*. Washington, DC: National Academies Press.

Bonilla-Silva, Eduardo. 2009. *Racism without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America*. 3 edition. Rowman & Littlefield Publishers.

Chen, Elsa, and Kevin Nomura. 2015. "And Justice for All? Racial and Ethnic Sentencing Disparities in California's Federal Drug Sentencing," May. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2611366.

Cooper, R., C. Rotimi, S. Ataman, D. McGee, B. Osotimehin, S. Kadiri, W. Muna, et al. 1997. "The Prevalence of Hypertension in Seven Populations of West African Origin." *American Journal of Public Health* 87 (2): 160–68.

DeSantis, Carol E., Rebecca L. Siegel, Ann Goding Sauer, Kimberly D. Miller, Stacey A. Fedewa, Kassandra I. Alcaraz, and Ahmedin Jemal. 2016. "Cancer Statistics for African Americans, 2016: Progress and Opportunities in Reducing Racial Disparities." *CA: A Cancer Journal for Clinicians* 66 (4): 290–308.

Dressler, William W., Kathryn S. Oths, and Clarence C. Gravlee. 2005. "RACE AND ETHNICITY IN PUBLIC HEALTH RESEARCH: Models to Explain Health Disparities." *http://dx.doi.org/10.1146/annurev.anthro.34.081804.120505*, September. https://doi.org/10.1146/annurev.anthro.34.081804.120505.

Foster, Thomas B., and Rachel Garshick Kleit. 2015. "The Changing Relationship Between Housing and Inequality, 1980–2010." *http://dx.doi.org/10.1080/10511482.2014.933118*, January. http://www.tandfonline.com/eprint/mF2r5kkvtNiXnmK5C45w/full?mobileUi=0.

Gee, Gilbert C., and Chandra L. Ford. 2011. "STRUCTURAL RACISM AND HEALTH INEQUITIES: Old Issues, New Directions." *Du Bois Review: Social Science Research on Race* 8 (1): 115–32.

Green, Nanny L. 1995. "Development of the Perceptions of Racism Scale." *Journal of Nursing Scholarship: An Official Publication of Sigma Theta Tau International Honor Society of Nursing / Sigma Theta Tau* 27 (2). Wiley Online Library: 141–46.

Hu, Li-tze, and Peter M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6 (1). Routledge: 1–55.

Jacoby, Sara F., Beidi Dong, Jessica H. Beard, Douglas J. Wiebe, and Christopher N. Morrison. 2017. "The Enduring Impact of Historical and Structural Racism on Urban Violence in Philadelphia." *Social Science & Medicine*, May. https://doi.org/10.1016/j.socscimed.2017.05.038.

Jones, C. P. 2000. "Levels of Racism: A Theoretic Framework and a Gardener's Tale." *American Journal of Public Health* 90 (8): 1212–15.

Khanna, Nikki. 2010. "'IF YOU'RE HALF BLACK, YOU'RE JUST BLACK':

Reflected Appraisals and the Persistence of the One-Drop Rule." *The Sociological Quarterly* 51 (1). Wiley Online Library: 96–121.
Krieger, N., J. T. Chen, B. Coull, P. D. Waterman, and J. Beckfield. 2013. "The Unique Impact of Abolition of Jim Crow Laws on Reducing Inequities in Infant Death Rates and Implications for Choice of Comparison Groups in Analyzing Societal Determinants of Health." *American Journal of Public Health* 103 (12): 2234–44.
Lichter, Daniel T., Domenico Parisi, and Michael C. Taquino. 2012. "The Geography of Exclusion: Race, Segregation, and Concentrated Poverty," August. https://doi.org/10.1525/sp.2012.59.3.364.
Link, B. G., and J. Phelan. 1995. "Social Conditions as Fundamental Causes of Disease." *Journal of Health and Social Behavior* Spec No: 80–94.
Lukachko, Alicia, Mark L. Hatzenbuehler, and Katherine M. Keyes. 2014. "Structural Racism and Myocardial Infarction in the United States." *Social Science & Medicine* 103 (February). Elsevier: 42–50.
Major, Brenda, Wendy J. Quinton, and Shannon K. McCoy. 2002. "Antecedents and Consequences of Attributions to Discrimination: Theoretical and Empirical Advances." In *Advances in Experimental Social Psychology*, 34:251–330. Academic Press.
Martin, David C., and Benjamin J. Newman. 2014. "Measuring Aggregate Social Capital Using Census Response Rates." *American Politics Research* 43 (4). SAGE Publications Inc: 625–42.
Massey, Douglas S., and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces; a Scientific Medium of Social Study and Interpretation* 67 (2). sf.oxfordjournals.org: 281–315.
Mayda, Anna Maria. 2006. "Who Is Against Immigration? A Cross-Country Investigation of Individual Attitudes toward Immigrants." *http://dx.doi.org/10.1162/rest.88.3.510*, October. https://doi.org/10.1162/rest.88.3.510.
Mendez, Dara D., Vijaya K. Hogan, and Jennifer Culhane. 2011. "Institutional Racism and Pregnancy Health: Using Home Mortgage Disclosure Act Data to Develop an Index for Mortgage Discrimination at the Community Level." *Public Health Reports* 126 Suppl 3 (September). JSTOR: 102–14.
Mendez, Dara D., Vijaya K. Hogan, and Jennifer F. Culhane. 2014. "Institutional Racism, Neighborhood Factors, Stress, and Preterm Birth." *Ethnicity & Health* 19 (5): 479–99.
Nei, M., and A. K. Roychoudhury. 1974. "Genic Variation within and between the Three Major Races of Man, Caucasoids, Negroids, and Mongoloids." *American Journal of Human Genetics* 26 (4). ncbi.nlm.nih.gov: 421–43.
Nunley, John M., Adam Pugh, Nicholas Romero, and R. Alan Seals. 2015. "Racial Discrimination in the Labor Market for Recent College Graduates: Evidence from a Field Experiment." *The B.E. Journal of Economic Analysis & Policy* 15 (3): 659.
Phelan, Jo C., and Bruce G. Link. 2015. "Is Racism a Fundamental Cause of Inequalities in Health?" *Annual Review of Sociology* 41 (1). Annual Reviews: 311–30.
Portes, Alejandro. 1998. "Social Capital: Its Origins and Applications in Modern

Sociology." *Annual Review of Sociology* 24 (1). Annual Reviews: 1–24.
Priest, Naomi, Yin Paradies, Brigid Trenerry, Mandy Truong, Saffron Karlsen, and Yvonne Kelly. 2013. "A Systematic Review of Studies Examining the Relationship between Reported Racism and Health and Wellbeing for Children and Young People." *Social Science & Medicine* 95 (October): 115–27.
Reardon, Sean F., John T. Yun, and Tamela Mcnulty Eitle. 2000. "The Changing Structure of School Segregation: Measurement and Evidence of Multiracial Metropolitan-Area School Segregation, 1989–1995." *Demography* 37 (3). Springer-Verlag: 351–64.
Reich, Michael. 1971. "The Economics of Racism." *Problems in Political Economy*. DC Heath Lexington, Mass., 107–12.
Sampson, R. J., S. W. Raudenbush, and F. Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277 (5328): 918–24.
Taylor, Teletia R., Thomas W. Kamarck, and Saul Shiffman. 2004. "Validation of the Detroit Area Study Discrimination Scale in a Community Sample of Older African American Adults: The Pittsburgh Healthy Heart Project." *International Journal of Behavioral Medicine* 11 (2). Springer: 88–94.
Utsey, Shawn O. 1999. "Development and Validation of a Short Form of the Index of Race-Related Stress (IRRS)-Brief Version." *Measurement and Evaluation in Counseling and Development* 32 (3). American Counseling Association: 149.
Wallace, Maeve, Joia Crear-Perry, Lisa Richardson, Meshawn Tarver, and Katherine Theall. 2017. "Separate and Unequal: Structural Racism and Infant Mortality in the US." *Health & Place* 45 (May): 140–44.
Williams, D. R., and R. Williams-Morris. 2000. "Racism and Mental Health: The African American Experience." *Ethnicity & Health* 5 (3-4): 243–68.
Wilson, William Julius. 1978. "The Declining Significance of Race." *Society* 15 (5). Springer: 11–11.
Zhou, Yuhong, Amin Bemanian, and Kirsten M. M. Beyer. 2017. "Housing Discrimination, Residential Racial Segregation, and Colorectal Cancer Survival in Southeastern Wisconsin." *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology* 26 (4). AACR: 561–68.

# Chapter 3

## Structural racism, race and body mass index in the United States

## Abstract

Over the course of three decades, the prevalence of obesity in the United States more than doubled, to 36% in 2010. Overweight and obesity are responsible for 5% of United States all-cause mortality and $147 billion in health care costs. Obesity prevalence is strongly patterned by race and gender, and the role of race in causing obesity has been an active area of research. Those efforts have been complicated by the absence of a precise definition of the biologic and social dynamics represented by the race variable, and also by the challenge of measuring structural racism. The current study is motivated by recent development of a validated scale measuring structural racism at the county level, and a conceptual framework that specifies hypothesized roles for race and structural racism in the development of adverse health consequences. We evaluate the association between county structural racism and BMI as reported in the 2011-2012 Behavioral Risk Factor Surveillance System, a large, national survey. We specified an interaction between county structural racism and black race on BMI. We found county structural racism was associated with lower BMI in both blacks and whites, although CSR was less protective in blacks than in whites. The results suggest further research is required to understand the appropriate geographic scale on which to measure structural racism, and to evaluate the possibility that gender, race and structural racism together are responsible for elevated BMI.

## Introduction

Over the course of three decades, the prevalence of obesity in the United States more than doubled, to 36% in 2010. (Fryar, Carroll, and Ogden 2016) Concurrently, the prevalence of extreme obesity, defined as body mass index greater than 40 $kg/m^2$, exploded, from 1.4% to 6.6% of the population. (Flegal, Panagiotou, and Graubard 2015) (2015) attribute 5% of United States all-cause mortality to overweight and obesity. Finkelstein, Trogdon, Cohen and Dietz (2009) estimate $147 billion in health care costs in the United States were attributable to obesity in 2008.

Race is a potentially important correlate of BMI and obesity. In 2012, obesity prevalence was 32.6 percent in non-Hispanic whites, and 47.8 percent in non-Hispanic blacks. (Ogden et al. 2014) Ailshire and House (2011) found BMI trajectories were ordered by social disadvantage, with educated white men having the lowest growth trajectory, and uneducated black women the highest. Burke et al. (1996) examined weight changes over five years in 4,207 young adults participating in the CARDIA study. They found black men gained weight at a higher rate than white men, and reported a similar result for black women.

Recently, investigators have called for additional research on discrimination and racism as a potential explanation for the racial disparity in many areas of health. (Bailey et al. 2017; Hicken et al. 2018) In the context of BMI, a major emphasis has been studies on BMI and interpersonal, or personally mediated, racism, defined as differential actions toward or assumptions about the abilities of, people according to their race. (Jones 2000) Another line of research has evaluated the link between BMI and structural racism,

defined as institutional policies and actions that result in detrimental treatment for members of a particular racial group. (Gee and Ford 2011) Discrimination and structural racism are constructs conceived of as involving five or more domains (Blank, Dabady, and Citro 2004), often including those of housing, education, health, criminal justice and employment.

Recent studies leave two important gaps in knowledge. To date, research on adiposity has most frequently analyzed individual domains of structural racism, rather than explicitly evaluating structural racism as an exposure. For example, Chang, Hillier, & Mehta (2009) reported a positive association between segregation and BMI among women, while Bower et al. (2015) found segregation was positively associated with odds of obesity among black women. Piontak and Schulman (2016) found a positive association between school segregation and childhood obesity BMI. Houle (2014) reported a similar finding with regard to male incarceration and BMI, noting as well a positive interaction between black race and incarceration on BMI. While literature on employment discrimination and BMI is sparse, Bhattacharya et al. (2004) found a positive association between poverty and BMI. Hernandez and Pressler (2014) reported similar findings with regard to the effect of childhood poverty on obesity in adolescence. The relationship between BMI and A1c control is more complex, in that excess BMI is often viewed as a risk factor for inadequate glycemic control (Nguyen et al. 2011; Hannon, Rao, and Arslanian 2005). Here, however, we specify a measurement model including county-level disparity in A1c, rather than A1c itself. The relationship between white/black A1c control ratio and BMI has not, to our knowledge, been studied

previously. Importantly, existing studies use single-indicator exposures, which are subject to both systematic bias and random error. The former can bias measures of association toward or away from the null; the latter most often underestimates the effect size. Investigators have examined the association between

One additional challenge is that investigators in this area have not specified a causal model that involves a modifiable intervention on race, along with a clearly specified set of covariates necessary for control of confounding. Several of the cited studies controlled for individual and neighborhood socioeconomic status, which may be powerful mediators of the effect of structural racism on health. Adjusting for potential mediators is likely to result in bias toward the null.

The current study is motivated by recent development of a validated scale measuring structural racism at the county level, and a conceptual framework that specifies hypothesized roles for race and structural racism in the development of adverse health consequences.

The conceptual model underlying the current study has been described in detail elsewhere. (Dougherty et al. 2018) Briefly, the model (Figure 1) decomposes the "race effect" into effects of structural racism, black (as opposed to white) physical phenotype, interpersonal racism, and cultural adaptation to structural racism. It posits that skin color and structural racism combine to affect access to resources, criminal justice involvement, and health. It specifies a sufficient cause interaction between race and CSR. In this relationship, subjects with black skin who are exposed to elevated levels of CSR exhibit higher BMI, while neither exposure on its own is sufficient to elevate BMI. In other

words, black skin is only relevant to BMI in settings of elevated structural racism, while structural racism is a risk for elevated BMI only if one is black. This specification follows from the literature emphasizing structural racism as a key determinant of health disparities (Krieger 2008), and the body of research demonstrating that phenotypic race is not a valid proxy for genotypic variation related to disease incidence. Specification of cultural adaptation to racism as a mediator of effect of structural racism on BMI is based on the work of Wilson (2009) and others, who have found many so-called cultural traits that are potentially detrimental to health are, in fact, responses to structural racism.

Development of the conceptual framework and measurement scale together offer the opportunity to examine previously untested causal relationships between structural racism, race and health. Development of the scale allows for more accurate measurement of structural racism, which in turn will yield enhanced estimates of the relationship between structural racism and BMI.

Previous analyses have investigated the effect of discrimination at the national, state, county, and area scale. We measure discrimination at the county level because it is the most relevant scale to the social dynamic under investigation. Racism often involves analytic units larger than the neighborhood. Housing segregation, for instance, often plays out in the context of municipalities or urban areas encompassing multiple municipalities. At the same time, action by institutions below the state level are relevant. County governments, for example, typically set educational policy that can result in school segregation. The results presented here may be subject to selection of a geographic scale.

## Methods

The procedure for estimating county-level structural racism is detailed elsewhere. (Dougherty et al. 2018) Briefly, we specified a latent variable measurement model with indicators spanning five domains of structural racism: employment, criminal justice, health, housing, and education (Table 1). For the first three domains, indicators were generally specified as prevalence ratios of the form $P_b/P_w$, where $P_b$ represents the proportion of a county's black residents experiencing a given phenomenon (unemployment, or residence in the county jail) among blacks, and $P_w$ represents the proportion of the county's white residents experiencing it. The final two domains included a mix of prevalence ratios and segregation indices commonly used in research on housing and education. (Massey and Denton 1988)

In order to limit random variability of the indicators, measurement was limited to U.S. counties with a black population of at least 500. Many indicators were derived from the 2011 five-year American Community Survey estimates provided by the U.S. Census Bureau. (2012) Additional sources of information include school demographic data compiled by the U.S. Department of Education, and county jail demographic data collected by the U.S. Department of Justice. The measurement model specified an independent error component for each indicator. The final model exhibited adequate fit. Pearson correlation coefficients between the indicators (Table 2) ranged from 0.009 (poverty ratio and A1c control ratio) to 0.48 (Housing dissimilarity index and high school graduation ratio).

## Data sources

Self-reported BMI and race data were obtained from the 2011 and 2012 Behavioral Risk Factor Surveillance System (BRFSS), a survey administered by the Centers for Disease Control and Prevention (CDC) in cooperation with state health agencies. The BRFSS sample is designed to yield state-level estimates of health behaviors included in the survey. Households and cellular telephones are called at random, and a randomly selected adult from each responding household is interviewed. (CDC 2013)

We elected to obtain BMI from BRFSS because we sought to include a large and representative selection of U.S. counties that are likely to encompass the full range of CSR and BMI levels. BRFSS is one of few data sources to meet this requirement. The combined 2011-12 dataset covers approximately 2,500 of the 3,143 counties in the country. The missing counties arise because of a CDC policy that suppresses records from counties with fewer than 11 respondents.

BRFSS respondents are asked to select one or more applicable races from the following list: White, Black or African American, Asian, Native Hawaiian or Other Pacific Islander, American Indian, Alaska Native, Other. Those who list multiple races are asked to identify their preferred race from among the multiple selections. Our analysis relies upon preferred race supplied by the respondent. We used BRFSS data from 2011 and 2012 because they are the most recent contiguous years in which the survey used compatible sampling weights. In 2011, the CDC changed the BRFSS sampling procedure to include respondents who do not own a landline. This required

altering the sampling weights in a way that made estimates from 2011 and subsequent years incomparable to those from 2010 and earlier years. The CDC restricted access to county identifiers for BRFSS respondents in 2013.

### Statistical analysis

The BRFSS sampling strategy yields state-level estimates for health characteristics addressed in the survey. By contrast, our target population was U.S adults residing in counties that were included in both BRFSS and in the CSR factor analysis. There is no readily available set of survey weights that would be relevant to this target population. Accordingly, we elected to treat the BRFSS data as a convenience sample of adults in the target population, and to analyze without survey weights.

To account for random error due to the small number of respondents in some counties, we specified multilevel linear regression models with a county random intercept. We controlled for age and sex, centering those variables on the grand mean because we were primarily interested in the effect of a county variable (CSR) while controlling for individual variables (Enders and Tofighi 2007) To evaluate the relationship between CSR and race, we specified models including a binary variable for black race, county proportion black to control for level two confounding, and a three-way interaction term for race, CSR and sex.

We hypothesized that structural racism would, after adjusting for individual-level confounders, be negatively associated with BMI in the general population, and that black race would be positively associated with BMI. Following the conceptual model introduced earlier, we hypothesized a sufficient cause interaction between CSR and race,

in which elevated CSR would be associated with lower BMI for whites, and higher BMI for blacks. We hypothesized that black females would have the highest expected BMI, followed by black males.

Our focus was on additive interaction, as our data were distributed appropriately for linear regression modeling. The standard test of statistical significance in a sufficient cause interaction requires an assumption of monotonicity (i.e., that structural racism is never protective). This assumption is not appropriate in the context of our data and conceptual model; thus we test for sufficient cause interaction under a more stringent test that does not assume monotonicity. We evaluate $p_{11}$-$p_{01}$-$p_{10}$>0, where $p_{11}$ is the beta coefficient for black subjects in counties with high CSR, $p_{01}$ is the beta coefficient for white subjects in counties with high CSR and $p_{10}$ is the beta coefficient for black subjects in counties with low CSR. (VanderWeele, Chen, and Ahsan 2011)

We specified five models: Model A, which regressed BMI against individual covariates; Model B, which added black race; Model C, which added county structural racism; Model D, which tested the sensitivity of the results in Model C to three potential macro-level confounders: percentage of county residents living in a rural area, median county income, and Census region; and Model E, which replaced grand mean centered values for race and sex with an individual binary variable and a county mean proportion and added three-way interaction term for individual race, sex, and CSR, as well as the component two-way interactions.

## Results

Our study population included 678,878 respondents who lived in a county

surveyed by BRFSS and provided valid BMI data (Table 2.3). Respondents were 39% female, 75.4% white non-Hispanic, 10.4% black non-Hispanic, and 14.2% other race or ethnicity. Respondents from roughly half of U.S. counties (1,563 of 3,142) were included in the analysis.

In Model A, we found female sex was associated with BMI decrease of 0.632 (0.67, 0.59) kg/m$^2$. Age was associated with a BMI increase of 0.061 (0.06, 0.06) kg/m$^2$. In Model B, black race was associated with a BMI increase of 2.624 (2.56, 2.69) kg/m$^2$, while the age and female coefficients moved slightly away from the null. In Model C, structural racism was associated with a BMI decrease of 0.3 (0.37, 0.23) kg/m$^2$. Coefficients for age, sex and black race were substantively unchanged. In Model D, we found an residence in the West Census region was associated with a BMI decrease of 0.348 kg/m$^2$, as compared to the Northeast region. Residence in the South region was associated with BMI increase of 0.134 (0.00,0.27) kg/m$^2$, while residence in the Midwest region was associated with an increase of 0.349 (0.21, 0.49) kg/m$^2$. An increase of $10,000 in county median income was associated with a BMI decrease of 0.26 (0.29, 0.23) kg/m$^2$. An increase of one point in the percentage of county residents residing in a rural area was associated with a BMI increase of 0.011 (0.01, 0.001) kg/m$^2$. In this model, the structural racism and black race coefficients moved modestly toward the null. increase of one standard deviation in CSR was associated with a decrease in BMI of -0.095 (-0.16,-0.03) kg/m$^2$. Black race was associated with a BMI increase of 2.587 (2.52,2.65). The female coefficient was -0.723 (-0.77,-0.68) kg/m$^2$. Model C, which specified no level 2 controls, resulted in estimates that were modestly further from the

null.

We found including a three-way interaction term for gender, race and CSR resulted in potentially important changes to the model (Table 3.4, Figure 3.2). Predicted BMI for male whites was largely unchanged across levels of CSR. BMI for white women dropped 0.4 $kg/m^2$ with a change in CSR from -1SD to 1 SD. For black males, a similar change in CSR resulted in a BMI increase of 0.4 $kg/m^2$. Black females reported the highest BMI across all levels of CSR, with BMI rising 0.1 $kg/m^2$ as CSR increased from -1 SD to 1 SD. The interaction term was statistically significant ($p<0.05$). Coefficients for the age and level 2 covariates other than CSR were substantively the same as in the other models.

## Discussion

Our study's strengths include use of a large, high-quality survey dataset that is broadly representative of the adult population of the United States. We used multilevel models to account for clustering within county, and to address small sample size in some counties. The current study is one of the first to evaluate the effects of black race and structural racism on a health outcome using a validated measurement scale for structural racism at the county level.

Our study demonstrates how investigators interested in structural racism might apply the newly developed conceptual framework and structural racism measurement model. Our hypotheses about the relationship of structural racism to BMI were largely validated, in that CSR was associated with lower BMI in the general population, but lower BMI for whites and higher BMI for blacks. At the lowest levels of CSR, the white-

black disparity in BMI disappeared for men, while increasing levels of CSR were associated with widening disparity. This is consistent with our expectation that modeling CSR would largely eliminate the effect of race on BMI. For women, however, this was not the case. Black women reported BMI consistently higher than other subjects at all levels of CSR. One explanation for these findings is that CSR has a more significant effect on black men through their exposure to the criminal justice system, which may subsequently affect employment and access to other resources.

Our results suggest that interventions to ameliorate structural racism may help black men to achieve healthier BMI. However, the finding that CSR is generally associated with lower BMI raises the question of how interventions might be structured so that they reduce CSR and black-white BMI disparities without increasing BMI in other populations. An additional concern is the dynamic involving black women, who appear to be less sensitive to changes in BMI. Further research is necessary to understand the significance of this finding, and to determine whether a similar dynamic applies to other health outcomes.

One potential limitation of the current study is reliance on self-reported BMI. However, self-reported BMI is generally adequately correlated with BMI from clinical measurements. (McAdams, Van Dam, and Hu 2007) (2007) evaluated the relationship between self-reported BMI and BMI measured during a clinical visit for National Health and Nutrition Examination Survey, finding correlation of .95 in white subjects and .93 in blacks. This finding is consistent with previous work in suggesting that self-reported BMI is a valid measure of true BMI, and that there is little risk of bias due to differential

measurement error by race.

A similar limitation involves our use of self-reported race. It is possible that those who do not exhibit a black physical phenotype have identified themselves as black, and that those who do have identified themselves as white. There are other approaches available to empirically characterize skin color. However, they are not often used in large national surveys, so relying on them would have diminished the external validity and power of our study. It is possible that BMI, or variables associated with BMI, such as social status, are also associated with error in reporting of race (Penner and Saperstein 2008), which could result in qualitative bias in our study findings. This possibility has not been carefully evaluated in the context of epidemiologic studies, so its potential impact on our results is unclear. We note that most epidemiologic studies in which race is used as an exposure or covariate share this limitation.

Another issue arises from the public health significance of BMI. Obesity in itself is not a meaningful endpoint; it is an indicator of metabolic processes that are associated with coronary artery disease, diabetes, and other illnesses. The value of our study rests on the degree to which BMI is associated with these endpoints, and whether those associations vary by race. McAdams and colleagues did find modest differences by race in correlation between self-reported BMI and biomarkers related to disease. For example, the correlation between self-reported BMI and HDL cholesterol was −0.53 for whites and -0.44 for blacks; they reported a similar discrepancy for C-reactive protein. (Stevens, McClain, and Truesdale 2008) (2008) found the correlation between BMI and disease risk was lower than those between sagittal abdominal diameter and waist circumference

and disease risk. However, differences were not large. (Sun et al. 2010) (2010) evaluated correlations between BMI and fat mass and percentage fat measures produced by dual-energy x-ray absorptiometry (DEXA), a more accurate measure of adiposity. Correlations (.78 for percent fat, .92 for fat mass) were identical for black subjects and white subjects.

This study requires a measure of adiposity that is widely available across a diverse selection of U.S. counties in order to evaluate the relationship between CSR and BMI across a full range of CSR levels. Other measures of adiposity, such as DEXA and waist circumference, are not available in datasets produced by large national surveys. The use of BMI, as compared to other adiposity indicators, may yield a small amount of imprecision in assessing the health risk of CSR. However, once the relationship between CSR and BMI is understood, further investigation can address the question of CSR and downstream outcomes.

An additional limitation of the current study is its cross-sectional nature. CSR measurements were taken at one point in time, which in effect assumes that the CSR of a subject's current county of residence is the CSR they have always been exposed to. In reality, subjects move between counties, and counties may experience changes in CSR over time. Further, our analysis assumes that the subject's most recent CSR exposure is the most relevant one to BMI. CSR may have lagged effects, and effects that vary over the life course, but we were unable to evaluate this with our data. We note that one limitation commonly cited in cross-sectional studies, that of reverse causation, is less of a concern here. It is unlikely that individual BMI is a cause of county structural racism, or

of the interaction between race and structural racism.

Finally, structural racism operates on multiple geographic levels (neighborhood, county, metropolitan area, and state, for example). Our results suggest that exploration of a multilevel measurement model (Muthén 1994) is appropriate. It is also possible that a portion of the BMI differential for black women represents the exposure to structural racism inherent to living in the United States. Measuring variation in structural racism between counties will not capture the full effect of structural racism if all U.S. residents are exposed to a baseline level of structural racism that impacts BMI (Rose 1985). It may be that cross-national analysis will provide a clearer picture of the effect of structural racism on BMI.

**Figure 3.1: Conceptual framework—structural racism, skin color and health**

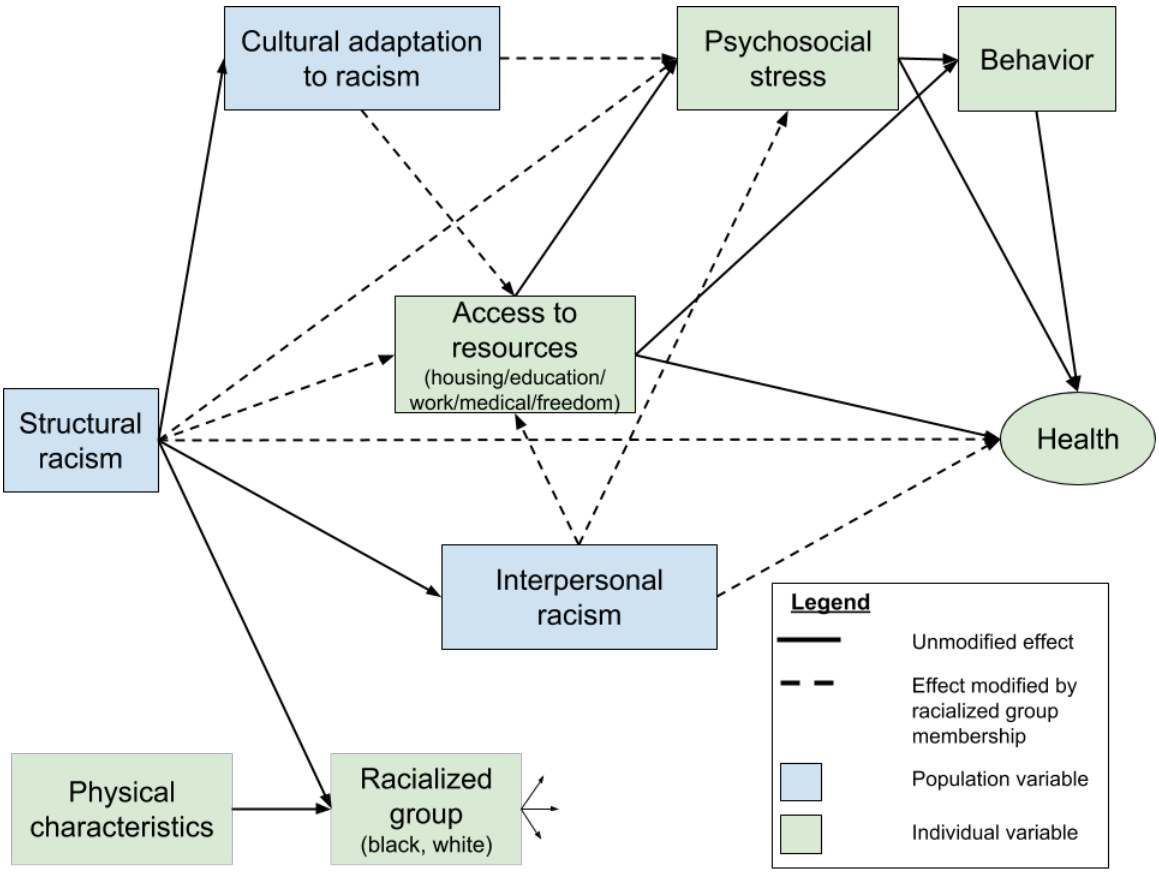


**Table 3.1: Items selected for structural racism factor model**

| Indicator | Domain | Data Source | Database | Description |
|---|---|---|---|---|
| Dissimilarity index | Housing | U.S. Census Bureau | American Community Survey SF-1 (2007-11) | Proportion of blacks that would have to relocate to achieve even distribution. |
| High-school graduation ratio | Education | U.S. Census Bureau | American Community Survey SF-1 (2007-11) | Ratio of non-Hispanic white to black high school graduation rates |
| Incarceration ratio | Criminal justice | U.S. Department of Justice | Census of Jail Inmates, 2005 | Ratio of non-Hispanic white to black county jail incarceration |
| Poverty ratio | Employment | U.S. Census Bureau | American Community Survey SF-1 (2007-11) | Ratio of non-Hispanic white to black poverty proportions |
| Diabetes prevention ratio | Health | Dartmouth Atlas of Health Care | 2012 Atlas | Ratio of white non-Hispanic diabetics receiving appropriate A1C monitoring to black diabetics receiving appropriate A1C monitoring |

**Table 3.2: Correlation matrix, confirmatory factor items**

| | Housing dissimilarity index | HS graduation ratio | Incarceration ratio | Poverty ratio | A1C control ratio |
|---|---|---|---|---|---|
| Housing dissimilarity index | 1 | | | | |
| HS graduation ratio | 0.475 | 1 | | | |
| Incarceration ratio | 0.279 | 0.152 | 1 | | |
| Poverty ratio | 0.149 | 0.185 | 0.119 | 1 | |
| A1C control ratio | 0.139 | 0.0899 | 0.0731 | 0.00923 | 1 |

**Table 3.3: Characteristics of included BRFSS respondents**

| | **Underweight** (<18.5 kg/m$^2$) | **Normal Weight** (<25 kg/m$^2$) | **Overweight** (<30 kg/m$^2$) | **Obese** (>=30 kg/m$^2$) | **p-value** |
|---|---|---|---|---|---|
| N | 11,844 | 232,794 | 244,399 | 189,841 | |
| Sex | | | | | <0.001 |
| Male | 78.7% | 67.6% | 50.4% | 58.7% | |
| Female | 21.3% | 32.4% | 49.6% | 41.3% | |
| Age, mean (SD) | 55 (22) | 54 (20) | 56 (17) | 55 (16) | <0.001 |
| Race | | | | | <0.001 |
| White | 77.3% | 79.3% | 76.5% | 71.0% | |
| Black | 6.9% | 7.0% | 9.8% | 15.5% | |
| Other | 15.8% | 13.7% | 13.7% | 13.6% | |
| Structural racism in county of residence, mean (SD) | -.0208 (.989) | .0258 (.992) | -.00199 (.996) | -.0331 (1.01) | <0.001 |

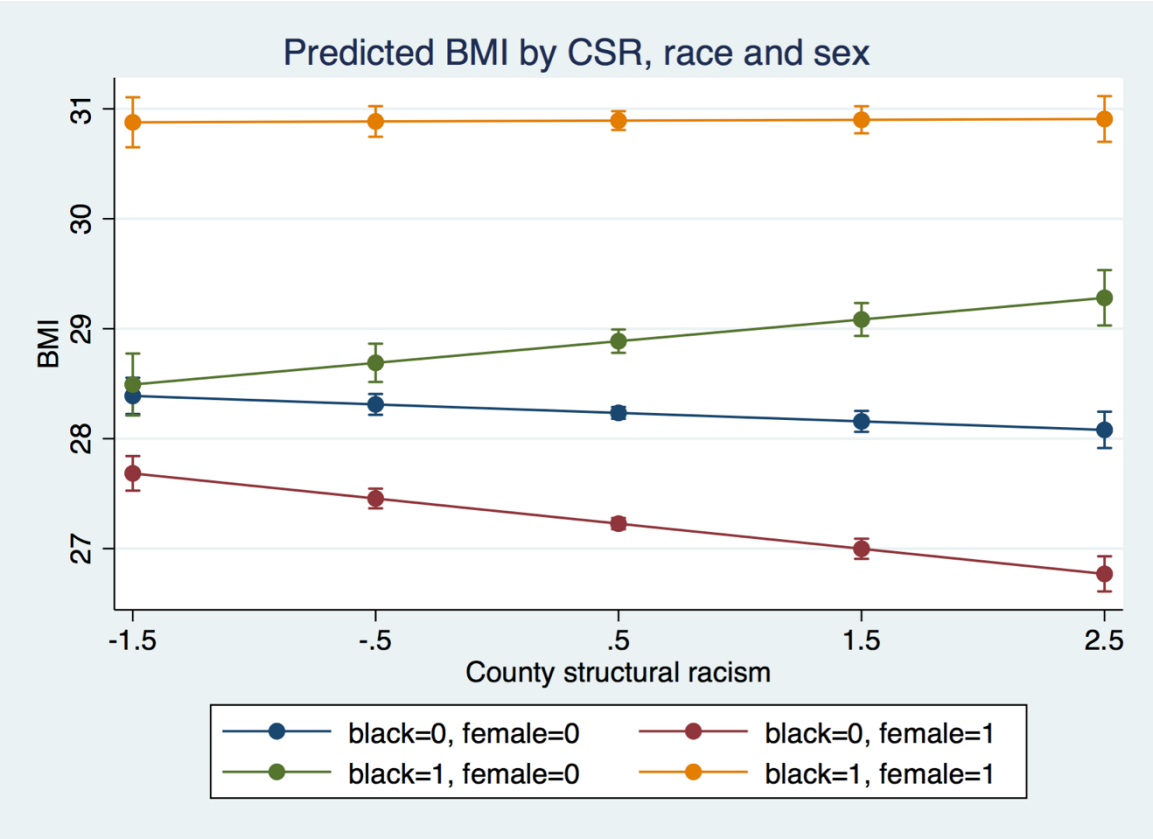
**Table 3.4: Regression coefficients, BMI on structural racism**

| | **Model A** | **Model B** | **Model C** | **Model D** |
|---|---|---|---|---|
| | Beta | Beta | Beta | Beta |
| | *[95% CI]* | *[95% CI]* | *[95% CI]* | *[95% CI]* |
| BMI | | | | |
| Female | -0.632 | -0.72 | -0.72 | -0.723 |
| | [-0.67,-0.59] | [-0.76,-0.68] | [-0.76,-0.68] | [-0.77,-0.68] |
| Age | 0.061 | 0.065 | 0.065 | 0.064 |
| | [0.06,0.06] | [0.06,0.07] | [0.06,0.07] | [0.06,0.07] |
| Black | | 2.624 | 2.633 | 2.587 |
| | | [2.56,2.69] | [2.57,2.70] | [2.52,2.65] |
| Racism | | | -0.3 | -0.095 |
| | | | [-0.37,-0.23] | [-0.16,-0.03] |
| Median income, $10k | | | | -0.26 |
| | | | | [-0.29,-0.23] |
| County percent rural | | | | 0.011 |
| | | | | [0.01,0.01] |
| Midwest | | | | 0.349 |
| | | | | [0.21,0.49] |
| South | | | | 0.134 |
| | | | | [0.00,0.27] |
| West | | | | -0.348 |
| | | | | [-0.51,-0.18] |
| Constant | 29.049 | 28.952 | 29.024 | 29.78 |
| | [28.98,29.11] | [28.89,29.01] | [28.97,29.08] | [29.51,30.05] |

**Table 3.5: Predicted BMI with CSR, race, sex interaction**

| | CSR=-1SD | CSR=1SD |
|---|---|---|
| White male | 28.3 | 28.2 |
| White female | 27.5 | 27.1 |
| Black male | 28.5 | 28.9 |
| Black female | 30.8 | 30.9 |

**Figure 3.2: Race, sex and CSR interaction**



## References


Ailshire, J. A., and J. S. House. 2011. "The Unequal Burden of Weight Gain: An Intersectional Approach to Understanding Social Disparities in BMI Trajectories from 1986 to 2001/2002." *Social Forces; a Scientific Medium of Social Study and Interpretation* 90 (2): 397–423.

Bailey, Zinzi D., Nancy Krieger, Madina Agénor, Jasmine Graves, Natalia Linos, and Mary T. Bassett. 2017. "Structural Racism and Health Inequities in the USA: Evidence and Interventions." *The Lancet* 389 (10077): 1453–63.

Bhattacharya, Jayanta, Janet Currie, and Steven Haider. 2004. "Poverty, Food Insecurity, and Nutritional Outcomes in Children and Adults." *Journal of Health Economics* 23 (4): 839–62.

Blank, Rebecca M., Marilyn Dabady, and Constance F. Citro. 2004. *Measuring Racial Discrimination*. Washington, DC: National Academies Press.

Bower, K. M., R. J. Thorpe Jr., G. Yenokyan, E. E. McGinty, L. Dubay, and D. J. Gaskin. 2015. "Racial Residential Segregation and Disparities in Obesity among Women." *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 92 (5): 843–52.

Bureau, Census. 2012. "2007-2011 ACS 5-Year Summary File Technical Documentation." U.S. Census Bureau. https://www2.census.gov/programs-surveys/acs/summary_file/2011/documentation/5_year/ACS_2007-2011_SF_Tech_Doc.pdf.

Burke, G. L., D. E. Bild, J. E. Hilner, A. R. Folsom, L. E. Wagenknecht, and S. Sidney. 1996. "Differences in Weight Gain in Relation to Race, Gender, Age and Education in Young Adults: The CARDIA Study. Coronary Artery Risk Development in Young Adults." *Ethnicity & Health* 1 (4): 327–35.

CDC. 2013. "Overview: BRFSS 2012." CDC Website. July 5, 2013. https://www.cdc.gov/brfss/annual_data/2012/pdf/overview_2012.pdf.

Chang, V. W., A. E. Hillier, and N. K. Mehta. 2009. "Neighborhood Racial Isolation, Disorder and Obesity." *Social Forces; a Scientific Medium of Social Study and Interpretation* 87 (4): 2063–92.

Dougherty, G. B., A. Gross, S. DeLuca, S. H. Golden, T. A. Glass, and L. T. Dean. 2018. "Development and Validation of a Scale Measuring County Structural Racism." *In Process*.

Enders, Craig K., and Davood Tofighi. 2007. "Centering Predictor Variables in Cross-Sectional Multilevel Models: A New Look at an Old Issue." *Psychological Methods* 12 (2): 121–38.

Finkelstein, Eric A., Justin G. Trogdon, Joel W. Cohen, and William Dietz. 2009. "Annual Medical Spending Attributable To Obesity: Payer-And Service-Specific Estimates," September. https://doi.org/10.1377/hlthaff.28.5.w822.

Flegal, K. M., O. A. Panagiotou, and B. I. Graubard. 2015. "Estimating Population Attributable Fractions to Quantify the Health Burden of Obesity." *Annals of Epidemiology* 25 (3): 201–7.

Fryar, C. D., M. D. Carroll, and C. L. Ogden. 2016. "Overweight, Obesity, and Extreme Obesity Among Adults 2009-2010." CDC/NCHS. 2016.

http://www.cdc.gov/nchs/data/hestat/obesity_adult_09_10/obesity_adult_09_10.htm.
Gee, Gilbert C., and Chandra L. Ford. 2011. “STRUCTURAL RACISM AND HEALTH INEQUITIES: Old Issues, New Directions.” *Du Bois Review: Social Science Research on Race* 8 (1): 115–32.
Hannon, Tamara S., Goutham Rao, and Silva A. Arslanian. 2005. “Childhood Obesity and Type 2 Diabetes Mellitus.” *Pediatrics* 116 (2): 473–80.
Hernandez, Daphne C., and Emily Pressler. 2014. “Accumulation of Childhood Poverty on Young Adult Overweight or Obese Status: Race/ethnicity and Gender Disparities.” *Journal of Epidemiology and Community Health* 68 (5): 478–84.
Hicken, Margaret T., Nicole Kravitz-Wirtz, Myles Durkee, and James S. Jackson. 2018. “Racial Inequalities in Health: Framing Future Research.” *Social Science & Medicine*, January. Elsevier. https://doi.org/10.1016/j.socscimed.2017.12.027.
Houle, Brian. 2014. “The Effect of Incarceration on Adult Male BMI Trajectories, United States, 1981-2006.” *Journal of Racial and Ethnic Health Disparities* 1 (1): 21–28.
Jones, C. P. 2000. “Levels of Racism: A Theoretic Framework and a Gardener’s Tale.” *American Journal of Public Health* 90 (8): 1212–15.
Krieger, N. 2008. “Does Racism Harm Health? Did Child Abuse Exist before 1962? On Explicit Questions, Critical Science, and Current Controversies: An Ecosocial Perspective.” *American Journal of Public Health* 98 (9 Suppl): S20–25.
Massey, Douglas S., and Nancy A. Denton. 1988. “The Dimensions of Residential Segregation.” *Social Forces; a Scientific Medium of Social Study and Interpretation* 67 (2). sf.oxfordjournals.org: 281–315.
McAdams, M. A., R. M. Van Dam, and F. B. Hu. 2007. “Comparison of Self-Reported and Measured BMI as Correlates of Disease Markers in US Adults.” *Obesity* 15 (1): 188–96.
Muthén, Bengt O. 1994. “Multilevel Covariance Structure Analysis.” *Sociological Methods & Research* 22 (3). SAGE Publications Inc: 376–98.
Nguyen, Ninh T., Xuan-Mai T. Nguyen, John Lane, and Ping Wang. 2011. “Relationship between Obesity and Diabetes in a US Adult Population: Findings from the National Health and Nutrition Examination Survey, 1999-2006.” *Obesity Surgery* 21 (3): 351–55.
Ogden, C. L., M. D. Carroll, B. K. Kit, and K. M. Flegal. 2014. “Prevalence of Childhood and Adult Obesity in the United States, 2011-2012.” *JAMA: The Journal of the American Medical Association* 311 (8): 806–14.
Penner, Andrew M., and Aliya Saperstein. 2008. “How Social Status Shapes Race.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (50). National Acad Sciences: 19628–30.
Piontak, Joy Rayanne, and Michael D. Schulman. 2016. “School Context Matters: The Impacts of Concentrated Poverty and Racial Segregation on Childhood Obesity.” *The Journal of School Health* 86 (12). Wiley Online Library: 864–72.
Rose, G. 1985. “Sick Individuals and Sick Populations.” *International Journal of Epidemiology* 14 (1): 32–38.
Stevens, J., J. E. McClain, and K. P. Truesdale. 2008. “Selection of Measures in Epidemiologic Studies of the Consequences of Obesity.” *International Journal of*

*Obesity* 32 Suppl 3: S60–66.
Sun, Q., R. M. van Dam, D. Spiegelman, S. B. Heymsfield, W. C. Willett, and F. B. Hu. 2010. "Comparison of Dual-Energy X-Ray Absorptiometric and Anthropometric Measures of Adiposity in Relation to Adiposity-Related Biologic Factors." *American Journal of Epidemiology* 172 (12): 1442–54.
VanderWeele, Tyler J., Yu Chen, and Habibul Ahsan. 2011. "Inference for Causal Interactions for Continuous Exposures under Dichotomization." *Biometrics* 67 (4). Wiley Online Library: 1414–21.
Wilson, William Julius. 2009. "Toward a Framework for Understanding Forces That Contribute to or Reinforce Racial Inequality." *Race and Social Problems* 1 (1). Springer US: 3–11.

# Conclusion

This dissertation advances the discussion surrounding race and health in a few ways: by developing a framework that identifies modifiable components of race and focuses interventions on modifiable upstream causes of health disparities; by identifying and validating an approach for measuring area structural racism; and by conducting the first empirical investigation into structural racism and a health outcome using the framework and measurement model.

The framework itself is perhaps the most important piece of this work. Although understanding the relationship between race and health has been a focus of epidemiology for a long time, the question of what we mean by race has not received nearly enough attention. At a minimum, I hope the dissertation demonstrates that race is complicated, and that any variable that comes with multiple definitions, and definitions that shift over time, is one scientists should treat carefully.

The process of thinking and writing about the theoretical underpinnings of race and health has emphasized for me the ease with which scientists, in many cases, jump into the fray without considering theory. There are some truly basic questions about race and health that not only remain unanswered, but which, until I began this research, had not been considered in the literature. If we describe race as unmodifiable, what is *Black Like Me* about? What type of effect do we believe structural racism has on white people, as opposed to blacks? What are the causal specifications of those dynamics? If we believe racism affects socioeconomic status, why do we adjust for SES when we're trying to measure the effect of racism? The framework provided in Chapter 1 purports to answer these questions, but in reality, it provides one set of what I suspect and hope will be many

collections of answers. The important thing is to begin asking the questions.

## Methodologic challenges

The measurement work detailed in Chapter 2 is significant because while many scientists believe structural racism is important, it has been poorly measured in the past. The factor model presented here is one way to remedy that problem, although again perhaps not the best and surely not the last. The measurement challenges in this area are formidable. As discussed, it is likely that everyone in the United States is exposed to a baseline level of structural racism that renders analysis of areal variation suspect. If one were to portion the variation in structural racism into between nation and within nation components, one suspects the within-unit variance would be the far less important component. This raises some difficult questions, because it means that the only way to get a robust measurement of the effect of structural racism is to conduct a cross-national analysis. However, the likelihood of non-comparable groups and off-support inference in such a design is high. Perhaps there is a subset of nations that have a large amount of variation in structural racism and are also comparable on some important covariates, and thus can serve as a study population for a cross-national look at structural racism.

The question of the appropriate level of measurement within the United States is also unanswered. It is possible that structural racism would be best measured at the level of state or metropolitan statistical area. More likely is that institutions on multiple geographic scales contribute to structural racism, and that measurement models accounting for this dynamic will be the most useful ones.

With regard to Chapter 3, the decision to apply the framework and measurement

model for the first time using BMI as an outcome had some drawbacks. The relationship between racism and BMI has not been as well studied as, for example, hypertension and infant mortality. Because the Chapter 3 analysis involved a new method for measuring the exposure and an exposure-outcome relationship that is not well defined, it is difficult to determine whether the findings reported here are related to mismeasurement of the exposure, or a dynamic specific to BMI. Had I selected another outcome on which to test the framework and measurement model, this would have been less of an issue. Using an outcome like infant mortality would have been advantageous for two additional reasons: It is not self-reported, and data exist for the entire U.S. population, rather than a sample.

The sampled data from BRFSS were a source of another challenge. BRFSS is not one sample; rather, each state designs a sampling procedure to yield state-level prevalence estimates. Because of this approach, states differ as to whether the primary sampling unit is county or some other geographic unit. This creates an issue, because generating appropriate level one and level two weights is critical to producing valid estimates from weighted sampling data. When the primary sampling unit is other than a county, the level two weights supplied by CDC aren't usable. A further challenge involved the question of target population. The BRFSS state samples could, with some adjustment to the weights, serve as an adequate study population for inference to the U.S. adult population as a whole in many instances. However, our Chapter 2 analysis produced structural racism factor scores for U.S. counties with black populations over 500. In addition, county identifiers were excluded from the BRFSS record for privacy reasons in small counties. So the study dataset contained exposure and outcome variables

for a subset of U.S. adults living in specific counties. In turn, our target population became the adult population of the counties included in the analytic dataset. There was no feasible way to adjust the CDC survey weights to produce estimates for this population, so I decided to treat the study dataset as a convenience sample and analyze without weights. Another approach would have been to proceed with generalized estimating equation modeling using survey weights raked to the population of the counties included in the study dataset. I decided to specify random intercept models because they yield more accurate estimates in smaller counties, where reported BMI may be subject to high levels of random error. However, it would be productive to review results from the GEE approach as a sensitivity analysis.

### Implications for future research

There are a number of possibilities for future research that emerge from the dissertation. The framework and measurement model provide a set of tools that can be used to evaluate the impact of structural racism on many types of health outcomes, as well as outcomes in other areas, such as education and housing. I have previously mentioned some avenues for methodologic research, including sensitivity analysis and exploration of additional specifications for structural racism measurement models.

Another question raised by the dissertation is the role of structural racism as a confounder. In many domains of research, particularly including health services research, analysts adjust for race as a confounder. One example of this is seen in studies comparing outcomes across a panel of hospitals, where the concern arises that race is a common cause of attendance at a given hospital and also of treatment outcomes. But if structural

racism, rather than race, is the determinant of hospital attendance and treatment outcomes, the strategies used to date to achieve control for confounding may result in substantial residual confounding. Exploring structural racism as a confounder may lead to more accurate results across a broad range of studies.

Finally, I hope the papers in this dissertation will spur research on structural racism and life-course epidemiology. Earlier, I mentioned the need to untangle the effects of parental vs. child exposure to structural racism. This is but one of many questions that require further analysis. Are there periods during human development that are particularly sensitive to structural racism? Do the effects of various types of limited access to resources differ across the life span? When will interventions on structural racism have the most impact? The framework and measurement model here may facilitate analyses addressing such questions.

## Biographical Statement

Geoff Dougherty is a doctoral candidate in the epidemiology department's general and methods track at Johns Hopkins Bloomberg School of Public Health. His research focuses on race, structural racism, and chronic disease outcomes. He works as director of health services research at U.S. News & World Report, leading a team that develops and implements quantitative methods for rating hospitals. Prior to transitioning to a career in public health, he spent 20 years as a journalist at leading U.S. news organizations, including the Chicago Tribune and Miami Herald. Dougherty holds a bachelor's degree in history and politics from The Colorado College, and a master of public health degree from Johns Hopkins. He was raised in Bethesda, Md., and now resides in Baltimore.