Title
Conceptualizing and Measuring Racism as a Multilevel Determinant of Health

Author
Michaels, Elizabeth Katahdin

Publication Date
2022

Peer reviewed|Thesis/dissertation

Conceptualizing and Measuring Racism as a Multilevel Determinant of Health

By

Elizabeth Katahdin Michaels

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Epidemiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Amani Allen, Chair
Associate Professor Mahasin Mujahid
Assistant Adjunct Professor Corinne Riddell
Chancellor's Professor Rucker Johnson

Spring 2022

Abstract

Conceptualizing and Measuring Racism as a Multilevel Determinant of Health

by

Elizabeth Katahdin Michaels

Doctor of Philosophy in Epidemiology

University of California, Berkeley

Professor Amani Allen, Chair

Racism is a fundamental cause of health inequities. Racism is multidimensional, encompassing: i) interpersonal (i.e., racial discrimination enacted by individuals); ii) institutional (i.e., discriminatory policies and practices occurring within social institutions); iii) structural (i.e., a web of inter-institutional connections which concentrate wealth, power, and health among whites); and iv) cultural (i.e., societal ideologies that place differential value on individuals based on race) processes. To effectively document racism, understand its effects on health, and develop interventions, it is necessary to measure it according to its dimension and the hypothesized pathways to health. This dissertation contributes to the growing racism and health literature, with a specific focus on the conceptualization and measurement of racism across its multiple distinct dimensions.

Chapter 1 examines the association between interpersonal-level racial discrimination and hypertension and depressive symptomatology, two health outcomes that are prevalent among African American women and increase risk of many leading causes of death, including cardiovascular disease. Data are from the African American Women's Heart and Health Study, a cross-section of African American women in the Bay Area with detailed survey and biomarker data. Given that racial discrimination is believed to harm health through repeated adaptation to chronic psychosocial stress, we investigate a novel approach to coding a commonly used discrimination scale to more accurately capture chronicity of racial discrimination experiences. Specifically, we develop and test a the "chronicity coding approach" and compare it to two conventional coding approaches. Findings suggest that scale coding influences exposure classification (i.e., low, moderate, or high levels of racial discrimination) and whether that discrimination is associated with hypertension but not depressive symptomatology. We discuss implications for scale coding in racism and health, and epidemiologic research more broadly.

Chapter 2 moves from the interpersonal- to the institutional-level. Using administrative data from the Home Mortgage Disclosure Act, we apply a novel measure of institutional racism at the census-tract level, namely the odds that a Black applicant is denied a home mortgage loan relative to an equally qualified White applicant, an indicator of institutional discrimination, neighborhood hostility, and resultant psychosocial stress. We link this measure with 2006-2015 data from the California Cancer Registry, a complete repository of all breast cancer cases in the state, with detailed information on tumor characteristics. Given increasing interest in the structural drivers of triple-negative breast cancer—an aggressive subtype most prevalent among Black females—we examine whether racial bias in home mortgage lending is associated with breast cancer incidence separately by race (non-Hispanic Black and non-Hispanic White) and subtype (triple-negative and Luminal A which is more common and has a more favorable prognosis). We find that racial bias in home mortgage lending is not associated with either subtype among either racial/ethnic group. Possible explanations for null findings, and directions for future research are discussed.

Finally, Chapter 3 presents a systematic review of the emerging literature on the health consequences of area-level racial prejudice, which I conceptualize as an indicator of cultural racism. Across fourteen studies reviewed, area-level racial prejudice is found to be associated with myriad adverse health outcomes, ranging from preterm birth to premature mortality. Findings are most pronounced among racial/ethnic minoritized groups, but several studies also find effects among Whites. After discussing conceptual and measurement considerations and illustrating potential pathways to health, we offer concrete directions for future research.

Taken together, these three chapters contribute to the growing literature on racism as a fundamental cause of health inequities. A primary theme of my dissertation is to contribute to the ongoing conversation of how we can best measure racism to understand its effects on health. I leverage data from a variety of sources—including a small community sample of African American women with rich social, psychosocial, and biomarker data; a large scale publicly available administrative dataset; a state-wide cancer registry; and the peer reviewed literature—to interrogate how racism operating across multiple dimensions is associated with a variety of mental and physical health outcomes and inequities. My dissertation confirms the adverse health consequences of racism, concluding that measurement decisions should be guided by the conceptual definition and level of racism, as well as the hypothesized social or biologic pathway to health. I hope this body of work will encourage the critical reflection and theory-driven measurement required to rigorously document and ultimately eliminate racism's harmful effects on the health of racially marginalized individuals, families, and communities.

This dissertation is dedicated to my Grandpa Sam,
who instilled in me the importance of integrity, family, and education,
and without whom none of this would have been possible. I love you.

ACKNOWLEDGEMENTS

As my mentor, Dr. Amani Allen always says, public health is a team science. I would not be where I am today without my dissertation committee and strong team of collaborators, consultants, and confidants. First and foremost, thank you to Amani, my beloved masters and PhD advisor, who has shaped me into the thinker I am today and models the kind of scholar and leader I someday hope to become. Amani has also served as a crucial source of social support through personal and professional highs and lows over the past seven years. I will be forever grateful for the time, care, and compassion she has showed me. Thank you to my other three outstanding dissertation committee members, Drs. Mahasin Mujahid, Corinne Riddell, and Rucker Johnson. Thank you to Mahasin for leading with compassion, providing pragmatic advice while uplifting my goals and dreams, and modeling humanity and humility. Thank you to Corinne for her critical feedback and hands-on support with big data wrangling, post-stratification weighting, and statistical modeling. I am grateful for her collegiate and collaborative mentoring style. Finally, thank you to Rucker for bringing a crucial interdisciplinary perspective to my work and training me third area: structural racism, social policies, and the lifecourse. I am grateful for his boundless generosity in sharing his data, time, and physical space with me this over the past several years.

In addition to my dissertation committee members, I am grateful to key mentors who have guided my scholarly development and provided support throughout my doctoral training: Drs. David Chae, Thu Nguyen, Scarlett Gomez, and Gilbert Gee. Thank you to my co-authors, Drs. Rebecca Hasson, Salma Shariff-Marco, Kirsten Beyer, Yehong Zhou, Rebecca Hasson, Melisa Price, and Christine Board for their collaboration on this work. I'd also like to thank Drs. Derek Griffith, Rudy Mendoza-Denton, Lonnie Snowden, Kirsten Bibbins-Domingo, and Len Syme for their helpful thought partnership throughout various stages of my training. Thanks to Drs. Maureen Lahiff, Suzanne Dufault, Nick Jewell, Jen Ahern, Alan Hubbard, Maya Petersen, and Patrick Bradshaw for offering methodological advice to strengthen the validity and rigor of my work.

Thank you to the three research groups, which together, served as my intellectual home during my graduate training: Dr. Allen's HEARTs Research Group, Dr. Mujahid's PLACE Research Group, and Dr. Nguyen's BD4HE Research Group. Having a safe and supportive space to share ideas, frustrations, and work in progress was critical to my success. I hope to someday create spaces as welcoming and collaborative as these for my students and mentees.

Thank you to the participants of the African American Women's Heart & Health Study, those individuals included in the California Cancer Registry, and all the participants in the studies included in my systematic review for sharing their data with the scientific community. Thank you to the authors in the systematic review for answering questions that arose along the way and for their important contributions to the field. Thank you also to the many undergraduate, graduate, and postdoctoral students working as

volunteer research assistants on the African American Women's Heart & Health Study and the California Cancer Registry. I am tremendously grateful to the Susan G. Komen TREND Fellowship Program and the National Heart, Lung, and Blood Institute for funding my dissertation research. Thank you to Dr. Art Reingold and the Epidemiology department for generously supplementing these training grants and enabling me to commit nearly 100% of my efforts to research. Thank you to Chao Guan, Dave Beza, and Raymond Fong for managing my grants, and to all of the BPH staff and administrators for all the incredible the work they do behind the scenes.

I would not have made it through this doctoral program without the unyielding support of my academic siblings and close confidants: Elleni Hailu, Rachel Berkowitz, Catherine Duarte, Alexis Reeves, Tracy Lam-Hine, Kevin Lee, Sean Darling-Hammond, Chris Lowenstein, Anoop Jain, Connor Martz, Xing Gao, and my Epi PhD cohort: Lauren Hunter, Isaac Ergas, Ruvani Jayaweera, Mary Horton, Moon Choi, Chris Rowe, Cam Adams, Sabrina Boyce, and Wendy Qi. Thank you for reading drafts of my work, answering my frantic coding and modeling questions, celebrating my wins with me, and commiserating with me during the tough times. Thank you especially to my fearless friend, collaborator, and forever role model, Dr. Marilyn Thomas, who demonstrated how to navigate this program with dedication, humor, and poise.

Finally, and most importantly, I wish to thank my family. Thank you to my incredible partner Max, who drove me to the mountains many weekends while I studied and wrote papers, who wiped away my tears during the hard times, and who proudly learned the meaning of a counterfactual. Thank you to my mom, who showed me that getting a PhD is possible and being a nerd is cool, who edited innumerable grad school applications, class assignments, and manuscripts, and will always help me with my lit reviews. Thank you to my dad for instilling a passion for health and medicine, and for always reminding me that even the best laid plans can and will go awry. Thanks to Bob, Chris, Mary Ellen, Tess, Robin, Grandpa Sam, and both Grandma Sues for their love and support. Lastly, thank you to my darling brothers, Walker and Jackson, for making me the person I am. I love you.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

LIST OF FIGURES

INTRODUCTION

Racism is a deeply-entrenched system of mutually reinforcing interpersonal, institutional, structural, and ideological processes operating overtly and covertly, implicitly and explicitly, to pattern the distribution of risk, resource, opportunity, and power on the basis of race.[1-10] In their landmark paper, Phelan and Link (2015) describe racism as a "fundamental cause" of racial health inequalities.[5] Briefly, fundamental causes are social conditions that (i) are associated with *multiple adverse health outcomes* via (ii) *multiple intermediary risk factors*.[11,12] These associations (iii) *persist over time*, despite *changing intermediary mechanisms (i.e., risk factors)*, and (iv) are characterized by access to *flexible resources (e.g., money, knowledge, power, prestige)* that can be utilized to avoid risk and lessen the impact of disease once it occurs.[11,13] Consequently, fundamental causes, such as racism, are most predictive of preventable health outcomes which are amenable to interventions involving flexible resources.[13,14]

Key examples of racial health inequities in the United States

As a fundamental cause that shapes the distribution of societal risks and resources, racism influences myriad health outcomes from preterm birth to premature mortality.[2,5-7,9,15-20] Two such health outcomes, which I examine in this dissertation, are cardiovascular disease (CVD) and breast cancer. CVD is the leading cause of death and largest contributor to the Black-White life expectancy gap in the United States.[21,22] According to recent estimates, 59% and 60% of non-Hispanic Black females and males suffer from CVD, compared to 42% and 54% among their respective White counterparts.[23] Proximal risk factors for CVD (e.g., hypertension) are important targets for prevention to improve population health and reduce health inequities.[24,25] Black persons have higher incidence of adverse cardiovascular outcomes than Whites across the lifecourse,[26-33] with disparities widening in middle-adulthood.[28,31,33] Further, the *rate* of health decline is differential by race, with Black adults showing equivalent cardiovascular risk profiles roughly 10 years earlier than Whites,[28,34-36] a pattern referred to as accelerated biologic aging, or *weathering.*[35,37]

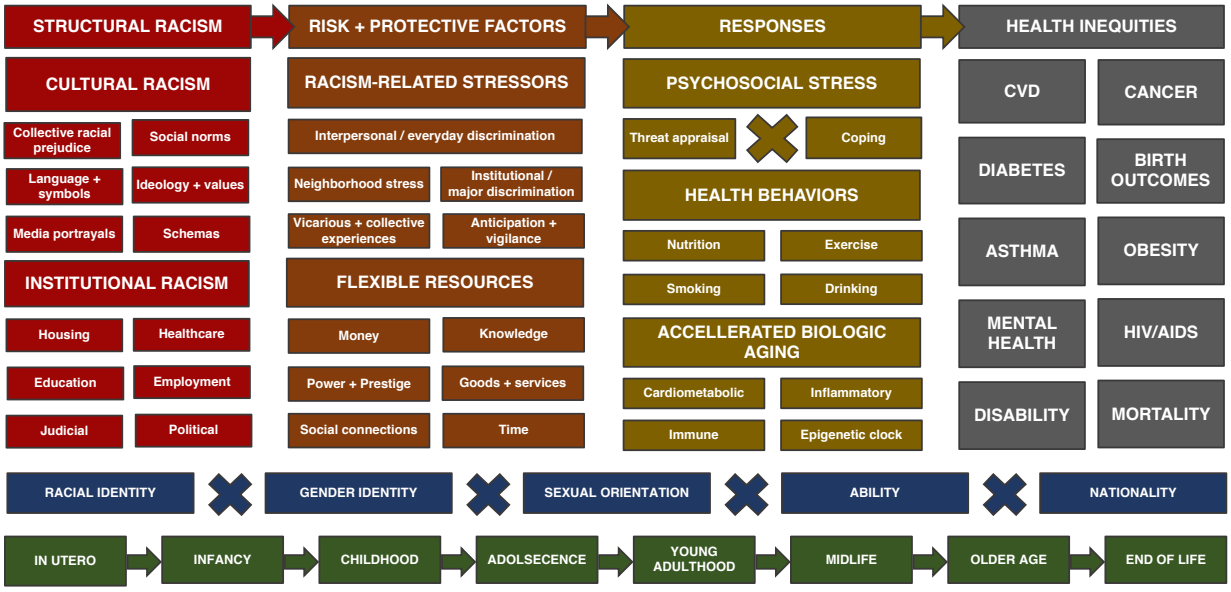
Breast cancer is the most commonly diagnosed cancer and the second-leading cause of death from cancer among females in the United States. Although breast cancer incidence is highest among NH White females, the burden of disease, including severity and mortality rates, is markedly higher among Black females. Moreover, although incidence rates have declined among most racial/ethnic groups, they have steadily increased among Black females, to a level that is similar to that of NH White females, the group with the historically highest rates.[38,39] Black females are also twice as likely as NH White females to be diagnosed with triple-negative breast cancer (TNBC),[40-43] which is more aggressive and less responsive to current treatments than the hormone receptor-positive (HR+) subtypes, such as Luminal A.[40,41,44] NH White females, in contrast, have the highest incidence of Luminal A breast cancer, the most

common subtype with the most favorable prognosis.[40,41] Once diagnosed, Black females have roughly 42% higher mortality from breast cancer compared to NH Whites.[45,46] Racial inequities have been documented across myriad other health outcomes, such as preterm birth and low birthweight,[47] diabetes,[48,49] low mental health,[50] and being injured or killed by law enforcement.[51,52]

Dimensions of racism

Racism, defined as "beliefs, attitudes, institutional arrangements, and acts that tend to denigrate individual or groups because of phenotypic characteristics or ethnic group affiliation,"[7] operates at multiple social levels, each of which impact health via distinct, yet reinforcing mechanisms.[1,6,7,16,53-56] While many taxonomies for racism have been proposed, we draw on frameworks offered by Jones (2000),[1] Williams, Lawrence, Davis (2019)[56] and Gee and Hicken (2021),[57] which together encompass interpersonal racism, institutional racism, structural racism, and cultural racism. Figure 1 displays my conceptual model, documenting these dimensions of racism and their pathways to health.

Figure 1. Conceptual model of racism dimensions and pathways to health



Interpersonal racism

Interpersonal racism, sometimes called "personally-mediated racism"[1] or "individual-level discrimination,"[56] refers to prejudice and discrimination enacted by individuals on the basis of race.[1,56] Prejudice describes differential assumptions about individuals based on their race, whereas discrimination refers to differential treatment on the basis of race.[1] Whereas racial discrimination can occur in everyday interactions or in institutional spaces, "interpersonal racism" refers to discrimination enacted between individuals in day-to-day life.

Everyday experiences of racial discrimination, such as being treated with less courtesy or respect because of one's race, are a pervasive racism-related stressor, commonly reported by Black Americans.[19,58] Chronic experiences of racial discrimination are hypothesized to degrade health through repeated affective and biological adaptation to chronic psychosocial stress.[7,19,59] Affective processes include threat appraisal and coping.[60] Biological processes include repeated activation of the body's physiologic stress response processes (e.g., hypothalamic-pituitary-adrenal axis), resulting in the over-circulation of stress hormones, such as cortisol.[61] Over time, this can lead to multisystem dysregulation, accelerated biologic aging (i.e., weathering[37]), and myriad chronic health conditions.[37,61]

Institutional racism

Racial discrimination also occurs in institutional settings. This form of discrimination can be described as institutional racism, or "racial inequity perpetuated by organizations, such as banks, hospitals, and governmental agencies."[56] Institutional racism occurs through formal/overt policies (e.g., Jim Crow Laws, redlining) as well as informal/covert practices, such as racial bias in policing and home mortgage lending discrimination.

Institutional racism harms health by shaping the distribution of societal risks (e.g., police, environmental pollutants, fast food and alcohol outlets, neighborhood stressors) and flexible health-promoting resources (e.g., education, income, healthcare, neighborhood green space, healthful foods) by race. Experiences of discrimination in institutional settings (e.g., being denied a bank loan, fired at work, or unlawfully stopped by the police) also act as racism-related stressors which, whether chronic or acute, can undermine health.[2,62]

Structural racism

Structural racism has been defined as "macrolevel systems, social forces, institutions, ideologies, and processes that interact with one another to generate and reinforce inequities among racial and ethnic groups"[55] and uphold "structural white supremacy."[57] Whereas some scholars use structural and institutional racism interchangeably,[56] this dissertation distinguishes these forms of racism, following Gee and Hicken (2021).[57] Specifically, institutional racism describes racially discriminatory policies and practices occurring within specific institutions (e.g., education or the judicial system), while structural racism is characterized by the "connections across multiple institutions, and the system as a whole."[57] For example, overt and covert policies and practices occurring in the education *and* judicial system act synergistically to maintain the school-to-prison pipeline and disproportionate incarceration of racially marginalized individuals.[63-65]

These inter-institutional connections are upheld by cultural processes, such as "racialized rules," defined as the "norms, principles, and regulations that govern the behavior of individuals and organizations that reinforce racial hierarchies."[57] Embedded in everyday practices, these racialized rules are normalized and deeply entrenched, such that they often operate invisibly to normalize inequality.[57]

Like institutional racism, structural racism concentrates flexible societal resources among whites, excludes racially marginalized groups from accessing these resources, and disproportionately exposes these communities to health-harming factors.[5,15,66] However, because it is maintained by inter-institutional connections, structural racism may produce health effects that are greater than the sum those caused by any particular institution (i.e., synergistic effects).[54]

Cultural racism

Cultural racism, defined as "the instillation of the ideology of inferiority in the values, language, imagery, symbols, and unstated assumptions of the larger society,"[56] creates a shared logical frame through which different racial and ethnic groups are valued.[56,67-69] As Williams et al. (2019) note, "cultural forms of racism may serve as the conduit through which views regarding the limitations, stereotypes, values, images, and ideologies associated with racial/ethnic minority groups are presented to society and are consciously or subconsciously adopted and normalized."[56] A related construct, advanced by Geronimus and colleagues (2016) is "the surround," or "subliminal reminders in our everyday rounds of the degree to which our social identity group is – or isn't – valued by society."[70] As a "pervasive logical structure," the surround guides our perceptions, expectations, and treatment of different identity groups, driving social and health inequies.[70]

Cultural racism is an important dimension of racism because it creates the ideological environment wherein institutional and interpersonal racism can flourish.[56] Building on Gee and Hicken's (2021) conceptualization of structural racism,[57] I argue that cultural racism serves as the ideological foundation upon which "racialized rules" are created and maintained. For example, assumptions about white as the dominant or superior race (i.e., culture of white supremacy) may undergird racialized rules around algorithmic decision-making, which create and maintain racial inequities across multiple social and health outcomes.[57] Thus, cultural racism may impact health via institutional racism and the resultant inequitable distribution of resources.[56,68,71] Cultural racism may also influence health by creating an environment that is more tolerant of interpersonal level racism,[56,69,72-74] leading to psychosocial stress and physiological dysregulation, as well as through vicarious racism and heightened vigilance in anticipation of institutional or interpersonal racial discrimination.[68,74-76]

Considerations for the measurement of racism in epidemiologic research

Accurate exposure assessment is fundamental to validity in epidemiologic research,[77] and research on the health consequences of racism is no exception.[78] Epidemiologic measurement decisions should be based on the conceptualization of the exposure and hypothesized biological mechanism linking the exposure and outcome.[77] Scholars have developed instruments and methodologies to measure racism in its various dimensions:

Much of the early work on racism and health relied on self-reported measures of interpersonal-level racial discrimination. One prominent example is the Everyday Discrimination Scale (applied in Chapter 1),[79] which asks survey respondents to recall experiences of discrimination in "day-to-day life", such as being followed around in stores, receiving poorer service in restaurants, etc. Respondents attributing these events to their race are classified as having experienced interpersonal-level racial discrimination.

Institutional racism can also be measured by self-report (e.g., "have you ever experienced discrimination when applying for a bank loan, seeking medical treatment, by the police or in the courts, etc.")[80] or by leveraging administrative data on home mortgage denials (as in Chapter 2),[81] employment discrimination, court sentencing decisions, healthcare referrals, or other forms of discrimination in institutional settings.

Structural racism is complex to measure empirically because it operates across rather than solely within social institutions.[9,57] Many studies aiming to measure structural racism may actually be measuring institutional racism because they focus on indicators of racism within a single institution as opposed to the synergistic interactions of discriminatory policies and practices operating across multiple institutions. Measuring structural racism requires creating indices that "reflect the multi-institutional nature of structural racism,"[57] including "formal and informal policies and practices, financial streams, and relationships occurring at the intersections of multiple societal institutions."[57]

Cultural racism is similarly challenging to measure, and there is currently no consensus or best practice; however, it can be argued that the increasingly popular approach of aggregating individual-level racial prejudice data to the area-level[71-73,82-90] (reviewed in Chapter 3) may provide a useful indicator of the "larger ideological environment,"[56] "unstated assumptions of larger society,"[56] and "shared national subconscious"[76] that are emblematic of cultural racism. Future research may consider other indicators of cultural racism, such as media portrayals or cultural symbols.

Thus, while racism is recognized as a fundamental cause of health inequities, there have until recently been a limited set of instruments to operationalize and quantify racism as a social exposure in epidemiologic research. Advances in big data, data harmonization, and interdisciplinary and cross-sector collaboration, have led to a

marked increase in the scope of data sources and methodologies used to measure racism as a multidimensional social construct.[78] As new data advances are developed and applied, it will be ever important to return to a strong theoretical conceptualization of racism, including the dimension of interest and hypothesized pathway to health.

This dissertation

The purpose of this dissertation was to measure racism at multiple social levels and examine associations with key health outcomes and health inequities in the United States. Of particular interest was how to best measure racism given the dimension of interest (e.g., interpersonal, institutional, cultural) and hypothesized pathway to health. Given the focus on racism as a fundamental cause of health inequities—which, by definition, influences multiple health outcomes via multiple pathways—I examined a variety of mental and physical health outcomes that are prevalent among Black Americans, a racial group that has faced disproportionate, egregious, and ongoing violence and discrimination in virtually every facet of U.S. society since the country's founding.

In addition to Fundamental Cause theory, this dissertation is guided by Nancy Krieger's ecosocial theory.[3,91] Ecosocial theory is grounded in four core constructs: (i) *embodiment*, or how we biologically incorporate our social and physical environments); (ii) *pathways to embodiment*, or the diverse mechanisms through which social and material exposures become biologically incorporated; (iii) *cumulative interplay* between exposure, susceptibility and resistance across the lifecourse, and (iv) *accountability and agency*, in terms of *who* is responsible for—and benefits from—the production and maintenance of social and health inequities.[3,91]

Chapter 1 was motivated by the tenant of *pathways to embodiment,* and more specifically, choosing a measurement strategy that most accurately quantifies chronic experiences of racial discrimination. Based on the understanding that racial discrimination acts as a source of chronic psychosocial stress which harms health through repeated activation of the body's stress-response processes, we challenged the dominant approaches to coding the Everyday Discrimination Scale, arguing that they may under-estimate the chronicity feature of racial discrimination. We developed a novel approach to coding the scale to capture chronicity more accurately. We then compared this coding approach to the two conventional coding approaches to examine potential differences in exposure classification and in associations with hypertension and depressive symptomatology among African American women in the Bay Area. We hypothesized that scale coding would influence exposure classification and associations with adverse health outcomes, and that the chronicity-based coding approach would yield the strongest associations due to more accurate assessment of chronicity, the etiologically relevant exposure.

Moving from the interpersonal-level to the institutional-level, Chapter 2 leveraged administrative data to measure racial bias in home mortgage lending and examined

associations with breast cancer incidence rate ratios among non-Hispanic Black and White females in California. By explicitly focusing on the institutional practice of racial discrimination in home mortgage lending, Chapter 2 contends with ecosocial theory's tenant of *accountability and agency.* Given increasing interest in understanding the structural drivers of triple-negative breast cancer[92]—a more aggressive and fatal subtype most prevalent among Black females[40]—we stratified analyses by subtype and race/ethnicity. We hypothesized that Black females living in communities with greater racial bias in home mortgage lending would be at increased risk of TNBC, whereas there would be no discernable effects among White females, based on differential exposure to social exclusion and associated psychosocial stress.

Chapter 3 provides a critical systematic review of the emergent literature using big data to measure cultural racism, with particular focus on conceptualization, measurement considerations, pathways to health, and directions for future research. We conducted a systematic literature review to describe evidence of the relationship between area-level racial prejudice and health, whether results differed by race/ethnicity, and to characterize key conceptual and methodological considerations to guide future research. We hypothesized that area-level racial prejudice would be associated with adverse health outcomes among multiple racial and ethnic groups, with stronger associations among marginalized racial/ethnic groups compared to non-Hispanic White populations. After reviewing and synthesizing the literature, we suggest directions for future research to further develop the conceptual and methodological rigor of this work and inform evidence-based interventions to advance population health and reduce racial health inequities.

CHAPTER 1: Coding the Everyday Discrimination Scale: Implications for Exposure Assessment and Associations with Hypertension and Depression Among a Cross-Section of Midlife African American Women

ABSTRACT

BACKGROUND: Studies suggest that racial discrimination impacts health via biologic dysregulation due to continual adaptation to chronic psychosocial stress. Therefore, quantifying *chronicity* is critical for operationalizing the relevant etiologic exposure and hence maximizing internal validity. Using one of the most common racial discrimination scales in the epidemiologic literature, we develop a novel approach for more accurately assessing chronicity and compare it to conventional approaches to determine whether coding influences differential exposure classification and associations with hypertension and depressive symptomatology among African American women.

METHODS: Data are from a socioeconomically diverse cross-section of 208 midlife African American women in Northern California (data collection: 2012-2013). Racial discrimination was assessed using the Everyday Discrimination Scale (Chronbach's α=0.95), and was coded using two conventional approaches: (1) situation-based coding: number of different situations ever experienced; (2) frequency-based coding: sum of Likert scale responses ranging from "never" to almost "everyday"; and (3) a new chronicity-based coding approach: sum of responses, weighted to capture annual chronicity (e.g., "a few times a month"=3x12=36x/year). Outcomes are hypertension and depressive symptomatology (10-item Center for Epidemiologic Studies Depression Scale (CES-D)).

FINDINGS: Exposure classification differed by scale coding approach, by up to 41%. There was a positive association between discrimination and hypertension prevalence for chronicity-coding only (Prevalence ratio=1.61, 95% Confidence interval=[1.03, 2.49]). For depressive symptoms, a dose-response relationship of similar magnitude was observed for all three coding approaches.

CONCLUSION: Scale coding is an important methodological consideration for valid exposure assessment in epidemiologic research. Coding can impact exposure classification and associations with important indicators of African American women's mental and physical health.

INTRODUCTION

African American women have higher rates of numerous adverse mental and physical health outcomes compared to other racial and gender groups[37,50,93-95] and a growing body of evidence implicates racial discrimination as a driver of these inequities.[6,7,37,79,96-98]

Racial discrimination, commonly reported by African American women, is hypothesized to impact health through repeated biological adaptation to *chronic* psychosocial stress.[7,19,59] Chronic, psychosocial stress can cause ongoing activation of the body's stress response processes, resulting in the over-circulation of stress hormones, which over time can lead to multisystem dysregulation and increased risk of poor health.[61] Quantifying *chronicity*, therefore, is critical for operationalizing the etiologically relevant exposure and hence maximizing validity among studies examining racial discrimination as a predictor of health.

The Everyday Discrimination Scale (EDS)—one of the most commonly used measures of discrimination in the epidemiologic literature[79,96]—is well-suited to measure chronicity. The EDS was developed as a measure to capture self-reported frequency of routine, relatively subtle discriminatory experiences in everyday social situations. First, respondents are asked: "In your day-to-day life, how often have any of the following things happened to you?" Examples include: "people treat you with less respect" and "people act as if they're afraid of you." Second, respondents identify the reason for the unfair treatment (e.g., gender, race, ethnicity).[79] Responses are typically coded on a 6-point Likert scale ranging from "never" to "almost everyday."[99-101]

While the EDS has been consistently and positively associated with adverse mental health outcomes among African Americans,[79,99,102,103] findings for physical health outcomes are less consistent.[96,97] For example, racial discrimination is often conceptualized as a risk factor for hypertension.[59,104,105] However, studies examining the association between racial discrimination—including studies using the EDS—and blood pressure outcomes show mixed results.[100,101,104-111]

One potential explanation for mixed findings across studies using the EDS is inconsistency in scale properties and the coding strategy used. These differences produce distinct measures of exposure, each with varying degrees of measurement bias, leading to differential associations with health. For example, the two most common approaches to coding the EDS—"situation-" and "frequency-based" coding[62]—differentially weight the survey response options, which may carry implications for assessing the *chronicity* of discrimination experiences.

In situation-based coding, each survey item is dichotomized: "never"=0, and "ever"=1.[62,102,106,109] As shown in Figure 1, situation-based coding collapses everyone who experienced *any* discrimination into one category (i.e., responses from "less than once a year" through "almost every day" are combined), obscuring the chronicity of

reports. Responses are summed across the items to generate a score ranging from 0-10, capturing the number of different situations *ever* experienced.

[FIGURE 1 HERE]

In frequency-based coding,[62,99-101,106-108] each response is given a value according to the Likert scale (never=1 to almost every day=6). Responses are summed across items to produce a score ranging from 10 to 60. Figure 2 illustrates that although frequency-based coding preserves distinctions between doses of the exposure, it assumes a monotonic change between each response. In reality, each successive response represents increasingly chronic experiences, which the frequency-based coding approach fails to capture.

[FIGURE 2 HERE]

In summary, the two common approaches to coding the EDS may underestimate the chronicity of everyday racial discrimination experiences. Given that racial discrimination is hypothesized to harm health via repeated physiologic adaptation to chronic psychosocial stress, underestimating chronicity may lead to misclassification of the etiologically relevant exposure. Such exposure misclassification may threaten internal validity and stall progress toward understanding the potential impact of chronic racial discrimination on racial health inequities.

Study aims

We develop a novel approach to coding the EDS, which scales each response to more accurately reflect the chronicity of discrimination experiences (Figure 3). We then compare our new coding scheme to the conventional situation- and frequency-based approaches to determine whether coding: (a) produces differential exposure classification and (b) influences the association of EDS with hypertension and depressive symptomatology among African American women.

Hypertension and depression are both stress-related conditions that disproportionately impact African American women, making them salient outcomes to examine in relation to racial discrimination within this population.[50,93] Moreover, these are two of the most commonly studied outcomes in the discrimination and health literature, which will allow us to compare our findings to existing work. We hypothesize that the association between the EDS and each study outcome will be differential based on coding approach; and that findings will be most pronounced using chronicity-based coding due to more accurate exposure assessment.

MATERIALS & METHODS

Study & recruitment

Data are from the African American Women's Heart Health Study, an exploratory, cross-sectional study examining associations between social and environmental stressors and mental and physical health among a community sample of 208 African American women in Northern California. We specifically recruited African American women to explore the unique health implications associated with navigating the intersection of multiple marginalized social identities (i.e., race *and* gender) in US society.[19]

Recruitment and data collection took place from March 2012 to March 2013. Study procedures are described in detail elsewhere.[112] Briefly, we utilized purposive sampling to maximize heterogeneity of sociodemographic factors and risk of experiencing racial discrimination. Participants were eligible if they self-identified as (1) African American, (2) female sex since birth, (3) aged 30-50, (4) US-born, (5) parent(s)/primary caregiver(s) are US-born African American, and (6) could read/write English. Exclusion criteria included: 91) pregnant or lactating, 92) self-reported a physician-diagnosed inflammatory or auto-immune disease. The study was approved by the Committee for the Protection of Human Subjects at the University of California, Berkeley.

Missing data

Systolic and diastolic blood pressure had the highest fraction of missing information (4.8%). We performed multiple imputation of missing values based on socioeconomic, psychosocial, and health status characteristics.[113] We excluded one respondent prior to imputation because data were missing for the majority of predictors (n=207). Relative variance increase was <10% for all models and relative efficiency was high (>98%).[114]

Study measures

Data collection included a computer assisted self-survey, in-person interview, and physical examination. Resting diastolic and systolic blood pressure was calculated as the average of three consecutive readings using an automated oscillometric monitor.[115] Hypertension (HT) was defined as: (a) systolic blood pressure (SBP) ≥ 140 mm Hg and/or, (b) diastolic blood pressure (DBP) ≥ 90 mm Hg, and/or (c) self-reported current cardiovascular medication usage.[116]

Depressive symptomatology was assessed using the 10-item Center for Epidemiologic Studies – Depression Scale (CES-D), which captures frequency of self-reported depressive symptoms in the past month (range: 0-30, α=0.83).[117-119]

Racial discrimination was assessed using the 10-item Everyday Discrimination Scale (EDS).[79] Because of the within-group design and focus on racially-based discrimination, we used a modified version of the EDS, which asks: "in your day-to-day life, how often have the following things happened to you *because of your race,*

ethnicity, or skin color?" Six response options range from "never" to "almost everyday" (Table S1).

We utilized three coding approaches to examine the potential effect of coding on exposure classification and associations with health outcomes (Table 1): (1) *situation-based coding* – we dichotomized each EDS item to "never"=0 and "ever" (collapsing those reporting "less than once a year" or greater into one category)=1. Items were summed (range: 0-10) to reflect the total number of situations "ever" experienced (α=0.89); (2) *frequency-based coding* – we scored responses according to the original Likert scale (range: "never" (1) to "almost everyday" (6)) and summed responses across items (range:10-60, α=0.95); (3) *chronicity-based coding* – we recoded each EDS response to reflect the total number of reported discrimination experiences, standardized upon the total number of days per year. "Never" was coded 0. "Less than once a year" was coded as the midpoint between 0 and 1 time per year=0.5x/year. "A few" is generally interpreted as 2-4, so we selected the midpoint=3. Therefore, we coded "a few times a year" as 3x/year and "a few times per month" as 3x12 months=36x/year. We coded "at least once a week" as 2x52 weeks=104x/year and "almost everyday" as 5x52 weeks=260x/year (Figure 3). Recoded items were summed to represent the total number of EDS experiences annually (range: 0-2600, α=0.95). To facilitate comparisons between coding approaches, we collapsed the three EDS measures into tertiles reflecting low, moderate, and high exposure (Table 1).[100]

[FIGURE 3 HERE]

[TABLE 1 HERE]

Covariate selection was outcome-specific and guided by directed acyclic graphs (Figure S1).[77] Confounders included: age, body mass index (BMI), neuroticism,[120,121] education, marital/partnership, and employment status. BMI was calculated as weight(kg)/height $(m)^2$. Age was confirmed via driver's license/state ID. All other covariates were self-reported. To increase parsimony, we modeled age and neuroticism continuously and dichotomized all other covariates.

Analysis

We generated Chronbach's alpha statistics and performed a polychoric principal components analysis (PCA) to evaluate internal consistency and scale dimensionality, respectively, under each EDS coding approach. All three coding iterations demonstrated a unidimensional data structure with high internal consistency, indicating recoding did not compromise the integrity of this previously validated scale. We used Goodman-Kruskal's Gamma-[122] and Cohen's Kappa[123] statistics to assess concordance/agreement in EDS exposure classification (low, moderate, high) between the three coding approaches, coded categorically. We also used Pearson's correlation coefficients to describe the strength of association between the continuous scales.

We fit multivariable modified Poisson regression models with robust standard errors to estimate hypertension prevalence ratios and 95% confidence intervals as a function of each EDS measure.[124] The prevalence ratio is a more appropriate measure of association than the odds ratio when the outcome is not rare (>10%), as was the case with hypertension in our sample.[124] Logistic regression models yielded similar results, albeit with the odds ratios over-estimating the PRs because the outcome was common (results not shown). Next, we fit multivariable linear regression models to estimate the association between each EDS measure and CES-D. All models used low EDS (bottom tertile) as the reference category. We report models unadjusted and controlling for covariates specific to the exposure-outcome relationship. All analyses were performed using Stata IC v13.[125]

RESULTS

Sample

Sample characteristics are displayed in Table 2. Systolic blood pressure was slightly elevated (mean=122, SD=20)), whereas diastolic blood pressure was in the normal range (mean=80, SD=12). Accordingly, 36% of the sample was hypertensive, less than the national prevalence of 46%.[93] The mean CES-D score was approximately 12 (SD=6), slightly above the recommended cut-off of 10 for depression.[118]

[TABLE 2 HERE]

EDS internal consistency and dimensionality by coding approach

Internal consistency was high for all three coding approaches (α=0.89, 0.95, and 0.95 for the situation-, frequency-, and chronicity-based coding approaches, respectively). Our PCA revealed a largely unidimensional data structure for all three approaches. For situation-based coding, the eigenvalue for the first component was 6.35 (63.5% of variance explained). The eigenvalue for the second component was 1.37. All other eigenvalues were <1. For both frequency- and chronicity-based coding, the eigenvalue for the first component was 7.57 (75.7% of variance explained). All other eigenvalues were <1.

EDS exposure classification by coding approach

Respondent distribution across low, moderate, and high EDS levels was differential by coding approach (Table S2). Table 3 summarizes the number (%) of respondents for whom exposure assessment was discordant (e.g., classified as low EDS using one coding approach and moderate or high using another), Pearson's correlation coefficients, Kappa statistics,[123] and Gamma statistics[122] for exposure classification agreement. The frequency- and chronicity-based approaches yielded the most concordant exposure classification and highest correlation/agreement, whereas the

situation- and chronicity-based approaches were most discordant and showed the lowest correlation/agreement.

[TABLE 3 HERE]

Associations between EDS and study outcomes by coding approach

Table 4 displays associations between each EDS measure and prevalence of hypertension. No association was observed using the situation-based coding approach. In contrast, we found an inverse U-shaped association using chronicity-based coding: moderate (versus low) levels of EDS were associated with a 61% higher estimated prevalence of hypertension (95% CI=1.04, 2.49), whereas high levels of EDS were associated with only a 10% increase in hypertension prevalence, and the 95% CI contained the null. We found a similar, but attenuated, U-shaped association between EDS and hypertension using the frequency-based approach: moderate levels were associated with a 10% higher prevalence of hypertension (95% CI contained the null), and associations of high levels were null.

[TABLE 4 HERE]

Table 5 displays results for depressive symptomatology. We observed a consistent dose-response association between EDS and CES-D, irrespective of coding approach.

[TABLE 5 HERE]

DISCUSSION

Accurate exposure assessment is fundamental to validity in epidemiologic research and measurement decisions should be based on the hypothesized biological mechanism linking the exposure and outcome.[77] We developed a novel approach to coding the Everyday Discrimination Scale for more accurately assessing the chronicity of everyday racial discrimination, a psychosocial stressor that is hypothesized to impact health via repeated stress adaptation. We compared our coding approach to conventional strategies to determine whether risk profiles and associations with health outcomes varied by coding approach. As hypothesized, the three coding schemes produced differential exposure classifications and associations with study outcomes. However, EDS coding was more instrumental for associations with hypertension than with depressive symptomatology.

Evidence of association between the EDS and hypertension is mixed;[100,101,106-110,126] our findings suggest this may be partially attributed to differences in coding approach, and hence, chronicity assessment. One prior cross-sectional analysis showed inconsistent associations between EDS and hypertension among African Americans when comparing situation- versus frequency-based coding strategies.[100] The present study

corroborates the finding of differential associations based on coding, and extends this work by developing and testing a novel chronicity-based coding approach.

That an association between EDS and hypertension was observed only using chronicity-based coding could be due to more accurate exposure assessment in relation to the proposed biologic pathways to health. Specifically, chronicity may be under-estimated by the frequency-based approach and entirely ignored by the situation-based approach. In both cases, non-differential exposure misclassification may bias results toward the null,[77] both here and in other studies seeking to measure chronic exposure to racial discrimination as a social determinant of health. The chronicity coding approach provides a more accurate exposure assessment, potentially reducing misclassification of chronic racial discrimination experiences. Findings may also help identify the mechanisms linking racial discrimination with cardiovascular functioning. Specifically, accurate chronicity measurement was crucial for modeling this association, which is consistent with stress theory and proposed biological pathway (i.e., repeated stress adaptation).[7,59,61,79,97]

The finding of higher hypertension prevalence among those reporting moderate, but not high, EDS parallels other work showing an inverse U-shaped relationship between racial discrimination and health among African Americans.[95,100,107,126] There are several plausible interpretations for this finding. First, appraisal and coping may differ by chronicity.[60,95,127] Those reporting moderate levels of EDS may possess fewer and/or less adaptive racism-specific coping strategies compared to those reporting high levels. Consequently, each encounter may be appraised as more stressful, resulting in exaggerated blood pressure reactivity, which over time may increase hypertension risk.[104] Second, acknowledging and reporting more chronic EDS may be indicative of a pro-black bias, greater race-centrality, and/or engagement in system- versus self-blame, all of which have been shown to buffer the effects of discrimination on health.[97,107,126,128] Finally, those reporting high EDS may have a blunted blood pressure response due to a lack of physiologic adaptation to chronic psychosocial stress.[61] Future research should explore these potential psychosocial and biologic mediators to further explicate the mechanisms through which EDS impacts health.

Unlike with hypertension, the association between EDS and depressive symptomatology was robust to coding. After adjusting for potential confounders, we saw a dose-response relationship using all three approaches. Findings are consistent with an extensive literature demonstrating a positive association between the EDS and depression, regardless of coding strategy used across studies.[99,102,103] Our study demonstrates that capturing chronicity is not critical for modeling the association between everyday racial discrimination and depression, suggesting there are other mechanisms at play that could be explored in future work.

Finally, results provide preliminary evidence of the construct and criterion validity for our new EDS coding scheme. Because frequency-based coding captures EDS with more granularity than situation-based coding, we would expect the chronicity-based

coding (the *most granular* assessment) to be more similar to frequency- than to situation- based coding both in terms of exposure classification and associations with study outcomes. Indeed, we found the highest correlation/agreement between chronicity- and frequency-based approaches (convergent validity) and the lowest correlation/agreement between chronicity- and situation-based coding approaches (discriminant validity). Relatedly, scale internal consistency and unidimensionality were more robust for the frequency and chronicity coding approaches than for the situation-based approach. This finding provides further evidence that the nature of the exposure may differ when items are dichotomized (situation-based approach) compared to when the gradient in experiences are retained (frequency- and chronicity-based approaches). The shape of association between EDS and hypertension was similar between the frequency- and chronicity-based coding approaches, while dissimilar from the situation-based approach (construct validity). While the frequency- and chronicity-based coding approaches were highly correlated and showed similar patterns of results, findings of association with hypertension were most robust using the chronicity-based approach, suggestive of criterion validity based on the proposed pathway to health. Thus, although the frequency- and chronicity-coding approaches are highly correlated, the latter provides a more nuanced exposure assessment, which may explain the more robust findings under this approach.

Next steps for future research include conducting a more formal exploratory and/or confirmatory factor analysis using these three coding approaches, testing these coding schemes on different health outcomes and more diverse study populations, and exploring whether sociodemographic (e.g., age, race, gender, socioeconomic status), coping (e.g., active vs. passive), or other psychosocial factors (e.g., social support) modify the scale validity, reliability, and health associations differentially by coding approach. Moreover, given previous evidence of differential validity and reliability of the EDS by race and by gender,[129,130] future research may also consider comparing psychometric properties of the three coding approaches stratified by these factors, a sub-analysis that was not possible in this within-group study of African American women.

This study had several limitations. Data are from a non-representative sample of African American women living in Northern California, and findings are not generalizable to the African American population as a whole. However, recruitment sought to maximize heterogeneity of exposure to discrimination and key covariates such as income and education, and characteristics of our sample are similar to the demographics of the source population, improving external validity.[131] The sociodemographic composition of our sample shares similarities with other national datasets; however, there are also important differences (e.g., participants in the present study had similar levels of education but lower prevalence of poverty compared to national samples of African Americans).[132-134] Previous work suggests that racial discrimination may manifest differently at various socioeconomic levels.[3,103,112] These findings should therefore be replicated in larger national samples to ensure

generalizability and interrogate whether the optimal coding approach differs based on the study sample's socioeconomic makeup.

While findings cannot be generalized to other gender or racial groups, the purpose of the study was to understand relationships between psychosocial stressors and mental/physical health among African American women, a particularly vulnerable group across numerous health indicators.[37,50,93-95] Additionally, the within-group study design uniquely facilitates an assessment of racial discrimination—rather than race—as the exposure of interest, a critical step toward understanding the drivers of racial health inequity.[135] Future research should explore these coding schemes in relation to health outcomes among African American men.

Cross-sectional data preclude causal inference; hypertension and depressive symptomatology could influence racial discrimination reporting. However, other studies have shown longitudinal associations between self-reported discrimination and incident hypertension[101] and depressive symptomatology,[102] indicating the potential directionality of these associations. Future research should apply these coding schemes to longitudinal data and test associations with disease progression. We also adjusted for neuroticism, a confounder of the association between discrimination and depression.[97,121] Neuroticism confounded the association between moderate EDS and depressive symptomatology for situation- and frequency-, but not chronicity-based, coding approaches (results not shown). Future research should explore whether reporting bias manifests differently depending on scale coding.

In developing the "chronicity" weighting structure, we made assumptions about the meaning each response (e.g., "a few" = 3). To assess potential measurement error, we conducted a sensitivity analysis under various assumptions (e.g., "a few" = 2 or 4). Results were largely unchanged, underscoring the robustness of this coding.

Chronicity coding is novel and cannot be directly compared to other studies. However, all three EDS iterations in our sample demonstrated high internal consistency and a unidimensional data structure, similar to other studies using the scale.[62,79,129,130] While distribution-based cutpoints are sample-specific, the tertile ranges shown in Table 1 facilitates reproducibility. Finally, although logistical constraints limited the sample to n=208, all models were powered >0.80.

CONCLUSIONS

Different approaches to coding the EDS produce distinct exposures that vary in their associations with important indicators of African American women's mental and physical health. Coding differences were more influential for associations with hypertension than with depressive symptomatology, which may help explain a puzzling pattern in the discrimination and health literature: consistent evidence for mental health outcomes, but inconsistent findings for physical, and particularly for cardiovascular, health outcomes.[96] Future research using the EDS should explicate hypothesized

mechanisms and code the scale accordingly. If the proposed pathway is through the *chronic* accumulation of discrimination experiences, then the chronicity-based coding approach may be most appropriate. This may be particularly relevant for strengthening internal validity in studies examining the association between racial discrimination and blood pressure outcomes among African American women.

TABLES AND FIGURES

Figure 1. Situation-based item weighting structure

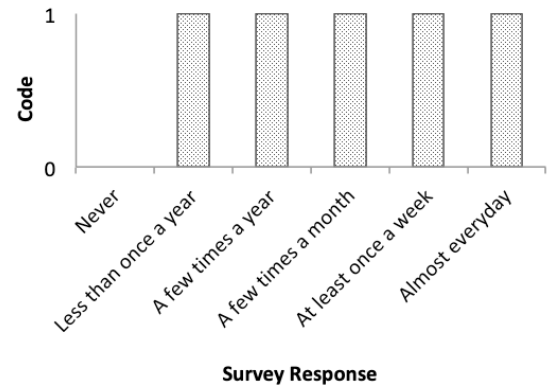


Figure 2. Frequency-based item weighting structure

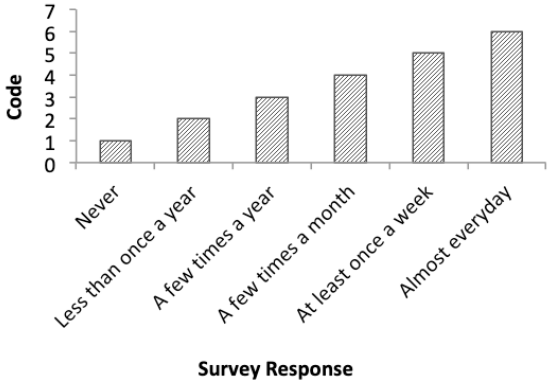


Figure 3. Chronicity-based item weighting structure

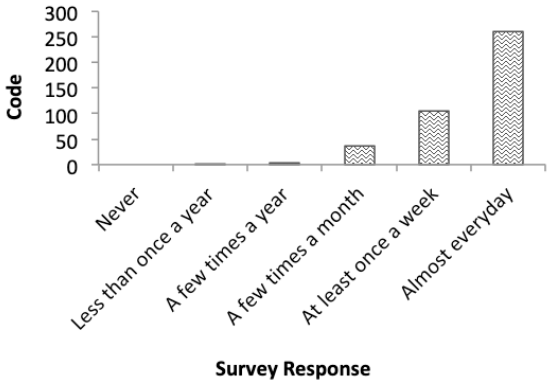


Table 1: Everyday Discrimination Scale survey response weighting structure, summary and tertile ranges, and description by coding approach, African American Women's Heart & Health Study, Northern California, 2012-2013 (n=207)

	Situation-Based Coding	Frequency-Based Coding	Chronicity-Based Coding
EDS Response			
Never	0	1	0
Less than once a year		2	0.5
A few times a year		3	3
A few times a month	1	4	36
At least once a week		5	104
Almost every day		6	260
Summary score range	0-10	10-60	0-2600
Lower tertile range	0-7	10-22	0-24
Middle tertile range	8-9	23-35	24.5-448
Upper tertile range	10	36-60	482-2600
Description	Number of situations "ever" experienced	Likert scale summary score	Total annual number of EDS experiences

Owing to the granularity of the chronicity coding, not all possible values of annual EDS experiences are represented.
EDS = Everyday Discrimination Scale.

Table 2: Sample characteristics, African American Women's Heart & Health Study, Northern California, 2012-2013 (n=207)

Variable	n (%) or mean (SD)
Poverty status, n (%)	
> 100% Federal poverty level	168 (81)
≤ 100% Federal poverty level	39 (19)
Educational attainment, n (%)	
> High school diploma	138 (67)
≤ High school diploma	69 (33)
Health insurance status, n (%)	
Insured	152 (73)
Not insured	55 (27)
Employment status, n (%)	
Employed	114 (55)
Not employed	93 (45)
Marital/partnership status, n (%)	
Married/partnered	61 (30)
Not married/partnered	146 (71)
Body mass index (BMI), n (%)[a]	
BMI ≥ 18.5 and <25	28 (14)
BMI < 18.5 or ≥25	179 (87)
Cardiovascular (CV) medication usage, n (%)	
Not currently taking CV meds	164 (79)
Currently taking CV meds	43 (21)
Age (years), mean (SD)	42 (5.9)
Neuroticism, mean (SD)[b]	3 (0.8)
Systolic blood pressure (SBP), mean (SD)[c]	122 (20)
Diastolic blood pressure (DBP), mean (SD)[c]	80 (12)
Clinically Diagnosable Hypertension, n (%)[c]	
Low on SBP and DBP and not taking CV meds	132 (64)
High on SBP or DBP or taking CV meds	75 (36)
CES-D Score, mean (SD)	12 (6.3)

[a] 7 cases (3.38%) missing
[b] 4 cases (1.93%) missing
[c] 10 cases (4.83%) missing
Due to rounding, some percentages do not sum to 100%.
CES-D = Center for Epidemiologic Studies-Depression.

Table 3. Discordance n (%), Pearson's correlation coefficient (r), Goodman-Kruskal's Gamma statistics (γ), and Cohen's Kappa statistics (κ) comparing Everyday Discrimination Scale exposure classification (low, moderate, high) by coding approach, African American Women's Heart & Health Study, Northern California, 2012-2013 (n=207)

Coding Approaches Compared	Discordance n (%)	Pearson's Correlation Coefficient (r)	Cohen's Kappa κ[a] (*P*)	Goodman-Kruskal's Gamma γ[b] (ASE[c])
Frequency vs. chronicity	34 (16)	0.90	0.75 (0.001)	0.97 (0.009)
Situation vs. frequency	64 (31)	0.76	0.54 (0.001)	0.88 (0.031)
Situation vs. chronicity	85 (41)	0.48	0.39 (0.001)	0.78 (0.044)

[a] Cohen's Kappa statistics test for agreement in exposure classification (low, mod, high) between EDS measures.[123]
[b] Goodman-Kruskal's Gamma statistics test for agreement in exposure classification (low, mod, high) between EDS measures, accounting for ordinal data structure with ties.[122]
[c] ASE = asymptotic standard error.

Table 4. Modified Poisson regression with robust error variances for association between racial discrimination and hypertension by Everyday Discrimination Scale coding approach, African American Women's Heart & Health Study, Northern California, 2012-2013 (n=207)

	Model 1: Number of EDS "Situations" Ever Experienced		Model 2: "Frequency" of EDS Experiences (Likert Summary Score)		Model 3: Annual "Chronicity" of EDS Experiences	
	Model 1a: Unadjusted	Model 1b: Fully Adjusted	Model 2a: Unadjusted	Model 2b: Fully Adjusted	Model 3a: Unadjusted	Model 3b: Fully Adjusted
EDS[a]	PR (95% CI)	PR (95% CI)	PR (95% CI)	PR (95% CI)	PR (95% CI)	PR (95% CI)
Moderate	0.97 (0.61, 1.56)	1.00 (0.63, 1.60)	1.14 (0.74, 1.76)	1.23 (0.81, 1.88)	1.52 (0.96, 2.39)	1.61 (1.04, 2.49)
High	1.12 (0.74, 1.70)	1.09 (0.72, 1.65)	1.01 (0.63, 1.61)	1.00 (0.63, 1.61)	1.14 (0.69, 1.90)	1.10 (0.66, 1.84)
Constant	0.36 (0.27, 0.48)	0.37 (0.20, 0.68)	0.35 (0.26, 0.48)	0.34 (0.18, 0.64)	0.30 (0.21, 0.44)	0.31 (0.17, 0.57)

Model a: Unadjusted.
Model b: Adjusts for age, body mass index, education (≤ HS diploma), marital/partnership status, and employment status.
[a] Referent group="low" EDS.
EDS = Everyday Discrimination Scale; PR = prevalence ratio; 95% CI = 95% confidence interval.

Table 5. Linear regression for association between racial discrimination and depressive symptomatology (CES-D) by Everyday Discrimination Scale coding approach, African American Women's Heart & Health Study, Northern California, 2012-2013 (n=207)

	Model 4: Number of EDS "Situations" Ever Experienced		Model 5: "Frequency" of EDS Experiences (Likert Summary Score)		Model 6: Annual "Chronicity" of EDS Experiences	
	Model 1a: Unadjusted	Model 1b: Fully Adjusted	Model 2a: Unadjusted	Model 2b: Fully Adjusted	Model 3a: Unadjusted	Model 3b: Fully Adjusted
EDS[a]	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Moderate	2.74 (0.65, 4.82)	1.16 (-0.69, 3.01)	3.15 (1.14, 5.17)	1.75 (-0.05, 3.54)	3.47 (1.45, 5.49)	2.01 (0.21, 3.81)
High	3.89 (1.90, 5.88)	2.37 (0.63, 4.11)	4.75 (2.72, 6.79)	3.03 (1.23, 4.83)	4.96 (2.93, 6.98)	2.85 (1.10, 4.69)
Constant	9.72 (8.40, 11.04)	8.75 (6.88, 10.61)	9.13 (7.72, 10.54)	8.39 (6.50, 10.28)	8.89 (7.45, 10.32)	8.34 (6.42, 10.25)
P for trend[b]	*P*=0.000	*P*=0.008	*P*=0.000	*P*=0.001	*P*=0.000	*P*=0.003

Model a: Unadjusted.
Model b: Adjusts for age, neuroticism, marital/partnership status, education (≤ HS diploma), and employment status).
[a] Referent group="low" EDS
[b] *P*-value (2-sided) associated with beta coefficient when EDS tertiles are modeled ordinally (versus categorically).
EDS = Everyday Discrimination Scale; 95% CI = 95% confidence interval.

CHAPTER 2: Home mortgage discrimination and incidence of triple-negative and Luminal A breast cancer among non-Hispanic Black and non-Hispanic White females in California, 2006-2015

ABSTRACT

BACKGROUND: In the United States, Black females are burdened by more aggressive subtypes and increased mortality from breast cancer compared to non-Hispanic (NH) White females. Institutional racism may contribute to these inequities. We aimed to characterize the association between home mortgage discrimination, a novel measure of institutional racism, and incidence of Luminal A and triple-negative breast cancer (TNBC) subtypes among NH Black and NH White females in California metropolitan areas.

METHODS: We merged data from the California Cancer Registry on females aged 20+ diagnosed with primary invasive breast cancer between 2006-2015 with a census tract-level index of home mortgage lending bias measuring the odds of mortgage loan denial for Black versus White applicants, generated from the 2007-2013 Home Mortgage Disclosure Act database. Poisson regression estimated cross-sectional associations of census tract-level racial bias in mortgage lending with race/ethnicity- and Luminal A and TNBC-specific incidence rate ratios, adjusting for neighborhood confounders.

RESULTS: We identified n=102,853 cases of Luminal A and n=15,528 cases of TNBC over the study period. Compared to NH Whites, NH Black females had higher rates of TNBC, lower reates of Luminal A breast cancer, and lived in census tracts with less racial bias in home mortgage lending. There was no evidence of association between neighborhood racial bias in mortgage lending at the time of diagnosis and either subtype among either racial/ethnic group.

CONCLUSION: Future research should incorporate residential history data with measures of institutional racism to improve estimation and inform policy interventions.

INTRODUCTION

Breast cancer is the most commonly diagnosed cancer and the second-leading cause of death from cancer among females in the United States. Considerable inequities in breast cancer severity and mortality have been documented between Black and non-Hispanic (NH) White females. Although Black females have historically shown lower incidence of breast cancer compared to their NH White counterparts, they are burdened by roughly 42% higher mortality from breast cancer once diagnosed.[45,46] Moreover, while incidence rates have declined among most racial/ethnic groups, they have steadily increased among Black females, to a national level that is now similar to that of NH White females.[38,39] Black females are also twice as likely as NH White females to be diagnosed with triple-negative breast cancer (TNBC), [40-43] which is more aggressive and less responsive to current treatments than the hormone receptor-positive (HR+) subtypes, such as Luminal A.[40,41,44] NH White females, in contrast, have the highest incidence of Luminal A breast cancer, the most common subtype with the most favorable prognosis.[40,41]

There is increasing recognition of the role of institutional racism, defined as "differential access to the goods, services, and opportunities of society by race"[1] in the production of racial inequities along the breast cancer continuum.[39,81,92,136,137] Housing discrimination is one primary form of institutional racism in the United States.[20,138-140] Although explicit discrimination has been illegal since the passage of the Fair Housing Act of 1968, covert forms of home mortgage discrimination against Black Americans persist, contributing to patterns of racial segregation that have persisted since the 1860s.[20,138-140] Emerging evidence documents associations between racial residential segregation and breast cancer incidence and mortality.[136,141-143] Importantly, however, racial residential segregation is a proxy measure that captures the consequence of institutional racial discrimination, not the discrimination itself.[144] In order to hold institutions and decision-makers accountable and inform policy change, there is a pressing need to rigorously interrogate the direct effects of discriminatory policies and practices, such as in home mortgage lending practices, on health.[3]

Publicly available data from the Home Mortgage Disclosure Act (HMDA) provide an opportunity for researchers to document the current extent housing discrimination across communities and examine its associations with health inequities.[53,81,144-146] The HMDA was enacted by congress in 1975 "to make lending practices transparent, ethical, responsive to community needs."[147] The Act requires lenders to report annual data on the location of housing loans and whether the loan was approved or denied, as well as demographic characteristics of the applicants.[147] Using these data, researchers can quantify racial bias in home mortgage lending, or the degree to which Black applicants are disproportionately denied loans relative to White applicants, adjusting for income and other relevant characteristics. [53,81,144-146]

Previous research has documented associations between home mortgage discrimination and health outcomes,[53,144,145] including breast cancer mortality.[81,146]

However, we are aware of no studies that have explored the relationship between housing discrimination and breast cancer incidence. This is an important topic because identifying risk factors for the development of cancer can inform primary prevention efforts. Neighborhoods where Black families face institutionalized discrimination and systemic exclusion could potentially represent toxic social environments which may be associated with greater breast cancer risk and the development of more aggressive subtypes.[92,148,149]

Breast cancer is a heterogeneous disease and subtypes defined by hormone-receptor (HR) biomarkers and human epidermal growth factor receptor (HER2Neu) status have distinct epidemiologic, etiologic, and prognostic profiles.[44,150-156] Given this, there is growing recognition among researchers and practitioners that breast cancer subtypes should be considered separate diseases.[150,152,155] Many commonly recognized behavioral and reproductive risk factors for breast cancer are associated with Luminal A and other HR+ subtypes, but not with the more deadly TNBC, which is more prevalent among Black females.[44,92,150-156] Identifying and intervening on potential structural drivers of the TNBC subtype is an urgent priority for achieving health equity.[92,151,154]

Therefore, we aimed to characterize the relationship between residence in neighborhoods with high home mortgage discrimination, a novel measure of institutional discrimination, and incidence of TNBC, and as a comparison, Luminal A breast cancer subtypes, among NH Black and NH White females in California. We estimated associations among NH Black and NH White females separately to examine how living in a community where there has been systematic exclusion of Black families from home ownership may differentially affect risk among these two racial/ethnic groups.

MATERIALS & METHODS

Data

We merged 2006-2015 breast cancer case data from the California Cancer Registry with census tract level measures based on 2007-2013 data from the Home Mortgage Disclosure Act database, 2007-2011 American Community Survey five-year estimates,[42] and population estimates based on the 2010 US Census.[157]

Breast Cancer Cases

Case data are from the California Cancer Registry (CCR, http://ccr.ca.gov/), a complete population-based repository containing detailed demographic, tumor, treatment, and survival information for all new cancer cases since 1988 (>3.5 million). The CCR comprises three of the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program registries (seer.cancer.gov/about). CCR data, derived primarily from the patient's medical record, included in this analysis were: age, race/ethnicity,

2010 census tract identifiers of residential address at diagnosis, and tumor HR (estrogen and progesterone receptor) and HER2Neu status, which were used to classify breast cancer subtype. Given the focus on understanding whether institutional racism is uniquely associated with triple negative breast cancer, the analysis was restricted to TNBC (HR-/HER2Neu-) and, as a negative control, Luminal A (HR+/HER2Neu-) subtypes.

All NH Black and NH White female cases of primary breast cancer aged 20+ and diagnosed in California between January 2006 and December 2015 were eligible for inclusion. We excluded cases of in-situ breast cancer, any cases that were not diagnosed with TNBC or Luminal A subtypes (including those for whom tumor subtype was not known), cases with a residential address unknown or not able to be geocoded, and cases residing outside of California metropolitan statistical areas (MSAs), as the racial bias index could only be calculated in metropolitan areas. This research was covered under the Greater Bay Area Cancer Registry (GBACR) IRB protocol # 18-24619 at the University of California, San Francisco. The GBACR IRB approval covers secondary analyses of de-identified cancer registry data without informed consent.

California Metropolitan Census Tracts

California census tracts with their centroid within a metropolitan statistical area were included in the analysis. Data at the census tract level included a racial bias index,[81] neighborhood stability, and neighborhood SES,[158] detailed below.

Racial Bias Index

Racial (anti-Black) bias in mortgage lending was estimated using 2007-2013 data from the Home Mortgage Disclosure Act (HMDA) database. Data were accessed from the Federal Financial Institutions Examination Council (FFIEC) HMDA website (http://www.ffiec.gov/hmda/).

We used a previously developed racial bias in mortgage lending index based on the HMDA data.[81] The index is described in detail elsewhere [81]. In brief, continuous surface maps were created for each MSA by interpolating the point estimates of the odds of home mortgage denial for a NH Black applicant as compared to a NH White applicant, adjusting for applicant sex and the income-to-loan ratio. These values were then averaged for each census tract to produce the "racial bias index," which was collapsed into quintiles for the analysis. Values > 1.0 indicate greater odds of denial for NH Black versus NH White applicants for the given census tract.

Covariates

Associations between racial bias in home mortgage lending and breast cancer incidence could be confounded by the socioeconomic resources of a community. Therefore, we included the Yang Neighborhood SES Index,[159] an adaptation of the Yost

Index for American Community Survey (ACS) data.[158] This is a composite of seven neighborhood indicator variables derived from the 2007-2011 ACS: (1) education index (median school years, percentage of high school graduates); (2) proportion with a blue collar job; (3) proportion older than 16 in the workforce without a job; (4) median household income; (5) proportion below 200% of the poverty level; (6) median rent; (7) median house value.[42,157,159] The neighborhood SES index was mean-standardized and included in models to adjust for confounding. As a measure of neighborhood stability, we adjusted for the percent of residents in a census tract who reported living in the same house one year ago on the 2007-2011 American Community Survey.[42] Given these neighborhood factors could also mediate the association between institutional racism and breast cancer incidence, we present adjusted and unadjusted models.

Statistical analysis

We used Poisson regression with generalized estimating equations (GEE, to account for clustering) to estimate population average incidence rate ratios (IRRs) and 95% confidence intervals (95% CIs) for associations between racial bias in mortgage lending and census tract-, race/ethnicity-, and age group-specific case counts, with the 2010 population count (times ten) as the off-set term.[157] We used an autoregressive (lag 1) [AR(1)] correlation structure to account for clustering by census tract. As a sensitivity check, we also ran Poisson models with exchangeable and independent correlation structures. Results were similar, so only the AR(1) results are reported. Model 1 adjusts for age (quadratic transformation of mid-point of each age group). Model 2 adjusted for hypothesized confounders (which could also be potential mediators), neighborhood SES (mean-standardized) and neighborhood stability (continuous). The estimation of the racial bias index was implemented in R and all other statistical analyses was performed using SAS v9.4.

RESULTS

Of the 8,057 census tracts in California, 7,836 with their centroid within an MSA were included in the analysis. Table 1 displays descriptive statistics for these census tracts. The racial bias index was skewed right and ranged from 0.3 to 86.2 with a median of 2.2.

[TABLE 1 HERE]

All NH Black and NH White female cases of primary breast cancer aged 20+ and diagnosed in California between January 2006 and December 2015 were eligible for inclusion (NH Black n=19,563, NH White n=185,082). After exclusions, the final sample consisted of n=11,063 NH Black and n=107,318 NH White cases of breast cancer over the study period.

Table 2 displays the characteristics of individuals included in the analysis, overall and disaggregated by race/ethnicity and subtype. Relative to NH Whites, NH Black females

with breast cancer had a higher proportion of TNBC (26.2% of NH Black vs 11.8% of NH White cases) and a lower proportion of Luminal A subtypes (73.8% of NH Black vs 88.2% of NH White cases). Among both racial/ethnic groups, those with TNBC were on average younger than those with Luminal A subtypes. Relative to NH White females, NH Black females also resided in census tracts with lower socioeconomic status, less neighborhood stability, and less racial bias in home mortgage lending; however, within-race/ethnicity differences in census tract characteristics between the two subtypes were minimal.

[TABLE 2 HERE]

Table 3 shows the results of the Poisson regression estimating the association between racial bias in mortgage lending and breast cancer incidence. Model 1 shows a modest and positive association between racial bias and rates of Luminal A breast cancer among NH White females (Q5 IRR=1.05, 95% CI=1.02, 1.07), which is attenuated by the inclusion of neighborhood confounders in Model 2 (Q5 IRR_{adj}=1.02, 95% CI=0.99, 1.04). No association was found between racial bias and incidence of Luminal A breast cancer among NH Black females (Q5 IRR_{adj}=0.92, 95% CI=0.85, 1.00), nor between racial bias and TNBC among either racial/ethnic group (NHB: Q5 IRR_{adj}=0.94, 95% CI=0.83, 1.07; NHW: Q5 IRR_{adj}=1.01, 95% CI=0.95, 1.07).

[TABLE 3 HERE]

DISCUSSION

This is the first study, to our knowledge, to explore whether racial discrimination in home mortgage lending—a novel measure of institutional racism—is associated with breast cancer incidence. We estimated incidence of TNBC and Luminal A separately, given these two subtypes are etiologically and epidemiologically distinct, and in response to calls to better understand the potential structural drivers of aggressive breast cancer subtypes (like TNBC) among Black females.[92,154] Consistent with prior literature, we found the TNBC subtype to be more prevalent among NH Black females with breast cancer, relative to NH Whites; and for both racial/ethnic groups, those presenting with TNBC were younger than those presenting with Luminal A breast cancer.[43] Our primary analysis revealed that census tract level racial bias in home mortgage lending, linked with breast cancer cases at the time of diagnosis, was not associated with incidence of Luminal A or TNBC subtypes among NH Black or NH White females in California. There are several explanations for the null associations we observed.

First, the degree of protective versus harmful features of environments with higher racial bias in home mortgage lending may neutralize risk for breast cancer. On the one hand, communities with more housing discrimination, and the individuals who live in them, may possess more structural advantages and health-promoting resources. This interpretation is consistent with our data showing that census tracts with more racial bias in mortgage lending had higher socioeconomic status (median neighborhood SES

in racial bias Q1= -0.3; in Q5 = 0.2). Previous studies using HMDA data showed protective associations of housing discrimination with breast cancer mortality and other adverse health outcomes.[53,145,146] For example, Collin et al. (2020) found that females with breast cancer who lived in Georgia census tracts with greater racial bias in home mortgage lending had improved breast cancer survival compared to those in lower bias tracts.[146] This finding aligns with studies showing protective associations between racial bias in home mortgage lending with birth outcomes[145] and general health status.[53] Communities with greater levels of racial bias in home mortgage lending may have a greater concentration of socioeconomic and health-promoting resources and opportunities due to systematic investment.[53,146] In addition, there may be a selection effect in which individuals who are able to live in these exclusionary communities may possess more socioeconomic resources and greater means to access improved health.[53,145]

On the other hand, communities with more housing discrimination may be characterized by toxic levels of psychosocial stress, in addition to other exclusionary policies and practices which could harm the health of all community-members and Black residents in particular.[81,137,144,146] This interpretation is consistent with several studies in the Milwaukee, Wisconsin area, which found that racial bias in home mortgage lending was associated with higher rates of all-cause mortality among Black individuals living with breast cancer[81] and colorectal cancer.[144] In short, communities with higher levels of racial bias in home mortgage lending may be characterized by both health-promoting resources, and health-damaging exposures; these conflicting forces may neutralize to produce the null associations we observed.

Another plausible explanation for our findings is that racial bias in mortgage lending measured at the time of diagnosis is not etiologically relevant for the development of breast cancer. Each of the prior studies that observed harmful or protective effects of housing discrimination on health examined outcomes which may be more sensitive to current social and environmental context, including mortality among those living with cancer,[81,144,146] birth outcomes,[145] and general health status.[53] Breast cancer has a long latency period with risk factors accumulating across the lifecourse and at key developmental stages.[160-162] Hence, linking neighborhood factors to address at the time of diagnosis limited our ability to measure the neighborhood context during more salient etiologic windows for the development of breast cancer, such as puberty.[160,161,163] While we adjusted for neighborhood residential stability in our models, the only data available was based on mobility from the past year, which does not account for all movement across the lives of individuals in our sample. Future longitudinal studies incorporating detailed residential history data with measures of racial bias and other forms of institutional racism could elucidate how structural conditions early in life are associated with breast cancer risk in adulthood, and inform targeted structural interventions during key etiologic windows.[161]
This study had several strengths. First, by leveraging data from the HMDA, we were able to measure one type of institutional racism directly, rather than the more common approach of measuring racial residential segregation as a proxy for discriminatory

practices.[144] The explicit examination of institutional racial discrimination, rather than its consequence, is important for illuminating injustice, increasing accountability, and informing policy interventions.[81,144,146] Second, compared to cancer registries in other states, the CCR is distinguished by collecting HER2Neu information starting in the mid-2000s, thus allowing for more granular subtype ascertainment over a longer period of time. Finally, our analytic design allowed us to estimate subtype-specific incidence of breast cancer. Several previous case-only studies demonstrated associations between racism-related exposures and odds of having a HR- versus HR+ subtype among females living with breast cancer.[149,164,165] However, by restricting the analysis to those diagnosed with breast cancer, the case-only design cannot estimate incidence.[166] In contrast, the use of a population denominator allowed us to directly estimate associations between census tract characteristics and incidence rate ratios for Luminal A versus TNBC subtypes, a novel contribution to the best of our knowledge.

This study also had several limitations. The measurement of census tract characteristics at the time of breast cancer diagnosis, rather than earlier in the lifecourse, may have limited utility for identifying neighborhood determinants of breast cancer risk. The analysis was restricted to California metropolitan statistical areas and results are not intended to generalize to more rural parts of California, nor to other states. In addition, census tracts are an imperfect proxy for neighborhoods and previous work found associations between neighborhood factors and cancer incidence differed by geographic unit of analysis.[142] The racial bias index does not capture discrimination against renters, potentially under-estimating the extent of institutional racism.[144,146] Moreover, the racial bias index was averaged from 2007 to 2013, which may mask changing dynamics of segregation and gentrification over the study period, which included the Great Recession and a period of increased housing foreclosures, both in California and nationally. Finally, while we conceptualized neighborhood socioeconomic status and residential mobility as potential confounders of the association between racial bias and breast cancer risk, they could also be on the causal pathway. We are unable to formally evaluate mediation using cross-sectional data; however, the similar pattern of results from our adjusted and unadjusted models suggests that any confounding or mediation by these factors was minimal. Relatedly, the CCR data lack potentially important individual-level socioeconomic, behavioral, and psychosocial factors which may mediate or moderate associations between neighborhood context and breast cancer incidence; exploring these mechanisms is an important direction for future research.

CONCLUSIONS

The explicit measurement of home mortgage discrimination and other forms of institutional racism in cancer research is a critical step toward identifying structural determinants, holding individuals and institutions accountable, and informing policy change. We applied a novel measure of institutional racism derived from publicly available data, which had previously been associated with survival outcomes among those living with cancer, to subtype-specific breast cancer incidence. Our null findings

of association between racial bias in home mortgage lending and incidence of Luminal A or TNBC may be due to neutralizing harmful and health-promoting forces, or due to the limitations and timing of the data. We caution against an interpretation of these findings that home mortgage discrimination does not increase breast cancer risk, but rather call for more research linking these and other measures of institutional racism in early life with the progression of breast cancer risk across the lifecourse.

TABLES AND FIGURES

Table 1. Characteristics of census tracts in California metropolitan statistical areas (n=7,836)

Census Tract Characteristics	Range	Mean (SD)	Median (IQR)
Racial bias index (n=7,836)[a]	0.3, 86.2	2.8 (2.6)	2.2 (1.5, 3.4)
Q1	0.3, 1.3	1.0 (0.2)	1.0 (0.8, 1.2)
Q2	1.3, 1.9	1.6 (0.2)	1.6 (1.5, 1.7)
Q3	1.9, 2.5	2.2 (0.2)	2.2 (2.0, 2.3)
Q4	2.5, 3.8	3.1 (0.4)	3.0 (2.7, 3.3)
Q5	3.8, 86.2	6.2 (4.1)	5.1 (4.3, 6.6)
SES index (n=7,775)[c]	-6.2, 3.1	0.0 (1.0)	0.0 (-0.8, 0.8)
Neighborhood stability (N=7,805)[d]	8.1, 100	84.2 (9.5)	85.9 (79.9, 90.6)

[a] Racial Bias Index=census-tract average odds of a mortgage denial for a Black applicant as compared to a White applicant, adjusting for individual sex, and the ratio of the loan amount to the applicant's gross annual income (data from Home Mortgage Disclosure Act, 2007-2013 (https://www.ffiec.gov/hmda/)).

[c] Yang Neighborhood SES Index=validated, census-tract level composite of median school years, percentage of high school graduates, proportion with a blue collar job, proportion older than 16 in the workforce without a job, median household income, proportion below 200% of the poverty level, median rent, and median house value, derived from 2007-2011 American Community Survey.[157,158]

[d] Neighborhood stability=census tract percent of residents who lived the same house one year ago on 2007-2011 American Community Survey.[157]

SES = socioeconomic status; SD = standard deviation; IQR = interquartile range.

Table 2. Age of primary breast cancer cases and characteristics of census tracts where they resided at time of diagnosis, by race/ethnicity and subtype, 2006-2015 (n=118,381)

	Overall			NH Black			NH White		
	Both subtypes (n=118,381)	Luminal A (n=102,853 (86.9%))	TNBC (n=15,528 (13.1%))	Both subtypes (n=11,063)	Luminal A (n=8,170 (73.8%))	TNBC (n=2,893 (26.2%))	Both subtypes (n=107,318)	Luminal A (n=94,683 (88.2%))	TNBC (n=12,635 (11.8%))
Age (median (IQR))	64 (54, 73)	64 (54, 74)	60 (51, 71)	60 (51, 70)	61 (52, 71)	58 (49, 67)	64 (54, 74)	65 (55, 74)	61 (51, 71)
Racial bias index (median (IQR)) [a]	2.2 (1.5, 3.4)	2.2 (1.5, 3.5)	2.2 (1.5, 3.4)	2.0 (1.4, 2.9)	2.0 (1.4, 2.9)	2.0 (1.4, 3.0)	2.3 (1.5, 3.5)	2.3 (1.5, 3.5)	2.2 (1.5, 3.5)
Q1 (%)	17.7	17.7	18.2	22.0	22.1	21.7	17.3	17.3	17.3
Q2 (%)	19.9	19.8	20.5	23.0	22.9	23.2	19.5	19.5	19.9
Q3 (%)	20.2	20.1	20.8	21.5	21.4	21.6	20.0	20.0	20.6
Q4 (%)	21.0	21.1	20.5	19.2	19.1	19.4	21.2	21.3	20.8
Q5 (%)	21.2	21.4	20.0	14.3	14.4	14.1	21.9	22.0	21.4
SES index (median (IQR)) [b]	0.5 (-0.2, 1.1)	0.5 (-0.2, 1.1)	0.3 (-0.4, 1.0)	-0.4 (-1.0, 0.3)	-0.4 (-1.0, 0.4)	-0.4 (-1.0, 0.3)	0.6 (-0.1, 1.2)	0.6 (-0.1, 1.2)	0.5 (-0.2, 1.1)
Neighborhood stability (median (IQR)) [c]	86.4 (81.0, 90.7)	86.5 (81.0, 90.8)	86.1 (80.5, 90.5)	85.7 (79.9, 90.3)	85.8 (80.0, 90.5)	85.4 (79.5, 90.1)	86.5 (81.1, 90.8)	86.6 (81.1, 90.8)	86.3 (80.8, 90.5)

Data restricted to invasive cases for whom census tract and tumor subtype information were known (Luminal A or TNBC), and who resided in California metropolitan statistical areas at time of diagnosis (n=118,381).

[a] Racial Bias Index=census-tract average odds of a mortgage denial for a Black applicant as compared to a White applicant, adjusting for individual sex, and the ratio of the loan amount to the applicant's gross annual income (data from Home Mortgage Disclosure Act, 2007-2013. (https://www.ffiec.gov/hmda/)); Q1: [0.3, 1.3), Q2: [1.3, 1.9), Q3: [1.9, 2.5), Q4: [2.5, 3.8), Q5: [3.8, 86.2].

[b] Yang Neighborhood SES Index=validated, census-tract level composite of median school years, percentage of high school graduates, proportion with a blue collar job, proportion older than 16 in the workforce without a job, median household income, proportion below 200% of the poverty level, median rent, and median house value, derived from 2007-2011 American Community Survey.[42,157,158]

[d] Neighborhood stability=census tract percent of residents who lived the same house one year ago on 2007-2011 American Community Survey.[42]

TNBC = triple negative breast cancer; SES = socioeconomic status; SD = standard deviation; IQR = interquartile range.

Table 3. Incidence rate ratios (IRRs) and 95% confidence intervals describing association between racial bias in mortgage lending and incidence of breast cancer in California metropolitan areas from 2006-2015, by race/ethnicity and subtype (n=118,381)

	Non-Hispanic Black		Non-Hispanic White	
	Luminal A (n=8,170)	TNBC (n=2,893)	Luminal A (n=94,683)	TNBC (n=12,635)
	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)	IRR (95% CI)
MODEL 1 [a]				
Racial bias index				
Q1	REF	REF	REF	REF
Q2	1.02 (0.94, 1.10)	1.03 (0.92, 1.15)	0.99 (0.96, 1.02)	1.01 (0.95, 1.07)
Q3	1.03 (0.95, 1.11)	1.04 (0.93, 1.16)	1.01 (0.98, 1.04)	1.03 (0.97, 1.09)
Q4	1.05 (0.97, 1.13)	1.08 (0.96, 1.21)	1.03 (1.00, 1.06)	1.00 (0.94, 1.06)
Q5	0.95 (0.88, 1.03)	0.95 (0.84, 1.08)	1.05 (1.02, 1.07)	1.01 (0.96, 1.07)
MODEL 2 [b]				
Racial bias index				
Q1	REF	REF	REF	REF
Q2	1.00 (0.93, 1.07)	1.02 (0.91, 1.15)	0.99 (0.96, 1.02)	1.01 (0.95, 1.07)
Q3	1.00 (0.93, 1.08)	1.03 (0.92, 1.16)	1.00 (0.97, 1.02)	1.03 (0.97, 1.09)
Q4	1.01 (0.93, 1.09)	1.07 (0.95, 1.20)	1.01 (0.98, 2.03)	0.99 (0.94, 1.05)
Q5	0.92 (0.85, 1.00)	0.94 (0.83, 1.07)	1.02 (0.99, 1.04)	1.01 (0.95, 1.07)

Data restricted to invasive cases for whom census tract and tumor subtype information were known (Luminal A or TNBC), and who resided in California metropolitan statistical areas at time of diagnosis (n=118,381).

Racial Bias Index=census-tract average odds of a mortgage denial for a Black applicant as compared to a White applicant, adjusting for individual sex, and the ratio of the loan amount to the applicant's gross annual income (data from Home Mortgage Disclosure Act, 2007-2013 (https://www.ffiec.gov/hmda/)); Q1: [0.3, 1.3), Q2: [1.3, 1.9), Q3: [1.9, 2.5), Q4: [2.5, 3.8), Q5: [3.8, 86.2].

[a] Model 1: Poisson regression model with census tract-, race/ethnicity-, and age group-specific case counts, with log of the population as the off-set term, and using generalized estimating equations with an AR(1) correlation structure with clustering by census tract, adjusted for age and age^2.

[b] Model 2: Same as Model 1, but also adjusted for the following census tract-level variables modeled continuously: Yang Neighborhood SES Index (validated, census-tract level composite of median school years, percentage of high school graduates, proportion with a blue collar job, proportion older than 16 in the workforce without a job, median household income, proportion below 200% of the poverty level, median rent, and median house value)[42,157,158] and neighborhood stability (census tract percent of residents who lived the same house one year ago on 2007-2011 American Community Survey) [42].

All census tract data are from the 2010 US Census,[157] 2007-2011 American Community Survey [42], and 2007-2013 Home Mortgage Disclosure Act (https://www.ffiec.gov/hmda/).

IRR = incidence rate ratio; 95% CI = 95% confidence interval; TNBC = triple negative breast cancer.

CHAPTER 3: Area-Level Racial Prejudice and Health: A Systematic Review

ABSTRACT

BACKGROUND: In recent years, there has been growing interest in “moving beyond the individual” to measure area-level racism as a social determinant of health. Much of this work has aggregated racial prejudice data collected at the individual-level to the area-level. Objective: As this is a rapidly emerging area of research, we conducted a systematic literature review to describe evidence of the relationship between area-level racial prejudice and health, whether results differed by race/ethnicity, and to characterize key conceptual and methodological considerations to guide future research.

METHODS: We searched four interdisciplinary databases for US-based, peer-reviewed articles measuring area level racial prejudice by aggregating individual-level indicators of racial prejudice and examining associations with mental or physical health outcome(s). Data extraction followed PRISMA guidelines and also included theory and conceptualization, pathways to health, and strengths and limitations.

RESULTS: Fourteen of 14,632 identified articles met inclusion criteria and were included in the review. Health outcomes spanned all-cause (n = 4) and cause-specific (n = 4) mortality, birth outcomes (n = 4), cardiovascular outcomes (n = 2), mental health (n = 1), and self-rated health (n = 1). All studies found a positive association between area-level racial prejudice and adverse health outcomes among racial/ethnic minoritized groups, with four studies also showing a similar association among Whites. Engagement with formal theory was limited and exposure conceptualization was mixed. Methodological considerations included unmeasured confounding and trade-offs between generalizability, self-censorship, and specificity of measurement.

CONCLUSIONS: Future research should continue to develop the conceptual and methodological rigor of this work and test hypotheses to inform evidence-based interventions to advance population health and reduce racial health inequities.

INTRODUCTION

Racism is a fundamental driver of health inequities.[1,2,5,9,56,167] Racism manifests in various forms, including, but not limited to, structural (i.e., institutional laws, policies, and practices), cultural (i.e., dominant ideological attitudes and beliefs about the relative value of different social groups) and personally-mediated (i.e., racial discrimination occurring in a variety of settings); each of which independently and interactively shape the distribution of societal risks, resources, and opportunities by race.[1,56]

While the majority of research to-date has focused on self-reported experiences of racial discrimination, there has been growing interest in "moving beyond the individual" to measure community-level racial prejudice as a social determinant of health.[75] One increasingly popular method involves measuring individual-level racial prejudice, aggregating to the area-level to capture the variation in racial prejudice across geographies, and examining associations with health outcomes and inequities.[71-73,82-90]

The rising popularity of this approach may be attributed to several factors, including: limitations of self-report for measuring exposure to racial discrimination, which can result in nonlinear and heterogeneous findings of associations with health outcomes;[96,112] a growing recognition of the role of structural and cultural racism in the production of health inequities;[9,15,56,70] interest in place-based determinants of racial differences in health;[168] and the rise of big data in social science and health research.[169,170] As this rapidly emerging area of research develops, a systematic and critical review of the evidence-to-date, including an examination of the conceptual foundations and methodological considerations, can help guide future research.

The purpose of this study was to systematically review all United States (US)-based, empirical, peer-reviewed studies examining the association between aggregated measures of area-level racial prejudice and health. Our primary research question was: what is the association between area-level racial prejudice and health outcomes in the US, and are associations differential by race/ethnicity? Secondary research questions were: (1) what theories or conceptual models frame the research?; (2) what are the hypothesized and empirically tested mediating and moderating pathways to health?; and (3) what are the primary measurement and other methodological challenges of this developing body of work?

DATA AND METHODS

All study procedures follow Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [171] and were preregistered: https://osf.io/vsze4/. Detailed methodology information, including search strings and data extraction codebook, can be found in Supplement 1.

Search strategy

We conducted an interdisciplinary search across the following electronic databases: (a) PubMed; (b) SCOPUS; (c) PsycInfo; and (d) Sociological Abstracts. Search terms were developed iteratively based on a preliminary review of the literature, research team expertise, and consultation with a public health research librarian.

Inclusion criteria were: (a) peer-reviewed journal article; (b) quantitative empirical study; (c) study conducted in the United States; (d) published in the English language; (e) study exposure is an indicator of bias, prejudice, animus, attitudes, sentiment, or beliefs toward or about a particular racial, ethnic, or immigrant group(s) that is measured at the individual-level and aggregated to the area-level; (f) study exposure is assessed using data from (1) surveys, questionnaires, or assessment tools, (2) social media, and/or (3) Google searches; and (g) study outcome is a mental or physical health outcome or health behavior. We utilized Covidence[172] for managing references.

For the purpose of this review, racial prejudice was defined as "differential assumptions about the abilities, motives, and intentions of others according to their race."[1] While the terminology used differs across the literature, we considered any measure of sentiment, attitude, belief, animus, prejudice, or cognitive (but not behavioral) bias toward or about specific racial, ethnic, and immigrant group(s) under the construct of "racial prejudice" when measured at the individual-level, and "area-level racial prejudice" when aggregated to the area-level. Note that prejudice is distinct from discrimination in that the former pertains to cognitive and affective aspects of racism whereas the latter describes its behavioral manifestation.[1] We excluded studies aggregating instances of interpersonal or institutional discrimination (e.g., hate crimes, housing or workplace discrimination) to keep the scope of the review narrowly focused on the cognitive and affective (i.e., prejudice) rather than the behavioral (i.e., discrimination) dimensions of racism. While much of the research in this area has focused on anti-Black racism, there is evidence documenting the unique manifestations, and harmful effects of, prejudice against various other racial and ethnic groups in the US.[173,174] Therefore, we included studies of aggregated prejudice against, and potential health consequences among, multiple racial/ethnic groups. We also included measures of anti-immigrant prejudice because people often conflate race/ethnicity with immigration status and immigrants often experience racism.[175]

We performed our database search on April 5, 2020. After removal of duplicates, 14,632 records proceeded to title and abstract screening, of which 32 advanced to full-text review, and 14 met inclusion criteria. Figure 1 shows the results of these exclusions.

[FIGURE 1 HERE]

Data extraction

Using MaxQDA,[176] we extracted data in accordance with PRISMA guidelines.[171] We also extracted data for our specific research questions. First, we were interested in conceptualization and framing: how were investigators thinking about area-level racial prejudice in relation to existing conceptual models for racism and health? Second, we extracted data on hypothesized and empirically tested mediation and moderation of the association between area-level racial prejudice and health outcomes. We were particularly interested in whether any association between area-level racial prejudice and health differed by race/ethnicity. Finally, we extracted data on key measurement and other methodological considerations.

All data extraction variables and the data extraction codebook were determined a-priori based on PRISMA guidelines and research questions. Two independent reviewers performed all screening and data-extraction, meeting regularly to build consensus. Any disagreements were resolved via consult with a third investigator. Our detailed methodology, including search strings and data extraction codebook, are shown in Supplement A.

RESULTS

Characteristics of included studies

Table 1 displays the characteristics of studies included in the systematic review. Ten studies were cross-sectional (i.e., exposure and outcome examined at one time-point only, even if the exposure preceded the outcome),[72-74,82-84,86,87,89,90] three studies examined outcomes prospectively using survival methods (i.e., exposure precedes outcome and there are multiple outcome assessments on each study participant),[71,85,88] and one employed a time-series analysis (i.e., changes in group-level rates over time).[177] All studies measured the exposure at the area-level, but the geographic scale ranged from the county (n=3)[86,87,90] to the national level (n=1),[177] with the largest number of studies (n=5) examining racial prejudice at the state-level.[71,73,74,84,89] Seven studies were multilevel, examining health outcomes at the individual-level using analytic methods that account for clustered data,[71,74,83,85,88-90] whereas one study did not account for clustering (i.e., individual-level study).[73] The remaining six studies were ecologic with the geographic area as the unit of analysis (e.g., rates as study outcome).[72,82,84,86,87,177]

Area-level racial prejudice was examined using one of four data sources: the GSS (n=3),[84,85,88] Project Implicit (n=4),[72,86,87,90] Google Trends (n=3),[71,82,83] or Twitter (n=4).[73,74,89,177] These data sources are described in detail below. Several studies specifically examined the racial bias of White (n=3)[72,86,87] and/or Black (n=2)[72,87] respondents. Four studies did not disaggregate the exposure by respondent

race,[84,85,88,90] and the remaining seven were unable to discern this information given the data available (i.e., Google or Twitter).[73,74,82,83,89,177]

Studies explored a variety of health outcomes, including birth outcomes (n=4),[74,83,89,90] all-cause mortality (n=4),[82,84,85,88] cause-specific mortality (n=4),[72,82,86,87] cardiovascular disease (CVD) and related risk factors (n=2),[73,86] mental health outcomes (n=1),[177] and self-rated health (n=1).[71] Six studies examined health outcomes of Black and White persons,[71,84-87,90] while four studies examined health outcomes of multiple (>2) racial/ethnic groups.[73,74,88,89] Three studies examined the health outcomes of Black persons only,[72,82,83] and one of Hispanics only.[177]

Ten of the fourteen studies explored whether associations between area-level racial prejudice and health was differential by racial/ethnic group, either by comparing race/ethnicity-specific rates[74,84,89] or by formally testing for multiplicative statistical interaction between area-level racial prejudice and race/ethnicity.[71,73,85-88,90] The remaining four studies were restricted to one racial/ethnic group.[72,82,83,177] Two studies explored mediators on the pathway from area-level racial prejudice to health outcomes.[85,86]

[TABLE 1 HERE]

Overview of data sources used to measure area-level racial prejudice

See Table S2 for more details of these data sources:

General Social Survey

The General Social Survey (GSS) is a nationally representative survey of social and political attitudes among non-institutionalized English-speaking adults living in the US, conducted on a random sample of the eligible population at each wave[178] (years available: 1972-2018). Examples of racial prejudice questions include: “Do Blacks tend to be unintelligent or tend to be intelligent?”, and “Do Blacks tend to be hard working or lazy?”.[85] Anti-immigrant prejudice is assessed based on agreement with statements like “America should take stronger measures to exclude illegal immigrants,” and “Immigrants increase crime rates.”[88] These questions are used to create composite measures of racial prejudice[84,85] or anti-immigrant prejudice,[88] which are then aggregated to the area-level.

Project Implicit

Project Implicit provides a free, online tool for assessing implicit and explicit biases toward various social groups (e.g., Black vs. White or gay vs. straight people) (years available: 2002-present).[179] Implicit racial bias is assessed via the “Implicit Association Test” (IAT), a timed dual-categorization task which measures the speed of keyboard associations between images of Black versus White faces and positive versus negative

words. Faster reaction time matching positive words with White and negative words with Black faces indicates cognitive dissonance between Black people and positive emotions, which is interpreted as an anti-Black and/or pro-White implicit bias.[72] Explicit bias is assessed via self-reported feelings of warmth or coldness toward Black/African American versus White/European American persons, and relative preference for the two groups.[72] Studies using Project Implicit data aggregate implicit and explicit biases separately to the area level, either for all respondents[90] or for subgroups of respondents (e.g., Whites' racial bias).[72,86,87]

Google Trends

Google search data come from Google Trends,[180] a publicly available dashboard where users can query the relative popularity of various search terms in specific geographic areas and time periods (years available: 2004-present). The data used in the included studies describe the popularity of Google searches for the "n-word" (ending in "-er(s)" but not "-a(s)") across 196 designated market areas (DMAs) from 2004-2007 (pooled). DMAs are geographic areas receiving similar media from television, radio, newspapers, and Internet sources.

Twitter

Twitter data can be collected prospectively[73,74,89] or retrospectively[177] (years available: 2006-present). Investigators can filter results to include tweets that reference particular racial/ethnic groups and use a combination of hand-coding, natural language processing, and machine learning to characterize the sentiment (positive, negative, neutral) of tweets referencing each group.[73,74,89,177] Using the latitude and longitude coordinates or other "place" information (e.g., city and state), investigators can geocode tweets to the area where they originated.

Summary of study findings

Measures of area level racism across multiple data sources were associated with a range of health outcomes, including CVD and related risk factors,[73,86] adverse birth outcomes,[74,83,89,90] increased mortality,[72,82,84-88] and poor self-reported mental[177] and physical health.[71] Although the overall pattern of results consistently showed a positive relationship between area-level racial prejudice and adverse health, the magnitude of association varied across studies (see Supplemental Table 1C). A direct comparison of effect sizes across studies is hindered by different exposure specifications (e.g., continuous vs categorical), measures of association estimated (e.g., beta coefficients vs odds ratios vs prevalence ratios, etc.), and other design features.

Evidence of differential health effects of racial prejudice by race/ethnicity was mixed: some studies found effects that were unique to, or more pronounced among, racial/ethnic minoritized groups; others showed no distinctions among racial/ethnic minoritized groups compared to Whites. In this review, we define the effects of area-

level racial prejudice on health to be differential by race/ethnicity if the p-value on the test for multiplicative interaction (i.e., race*area-level racial prejudice) was <0.05, following the threshold used in the reviewed studies and authors' interpretations of their results. If no test for interaction was performed, we considered the effects to be differential if the race/ethnicity-specific measure of association differed by >10%, a rule-of-thumb for a meaningful difference in epidemiologic quantities.[181]

Findings of Differential Effects by Race/Ethnicity

Among the ten studies that explored whether effects were differential by race/ethnicity, six found evidence of effect modification.

Of these, two showed harmful but less pronounced effects among Whites compared to racial/ethnic minoritized groups.[84,86] First, Kennedy et al. (1997) found that states with higher versus lower "collective disrespect" (based on GSS data) had higher Black and White mortality rates; however, these relationships were stronger for Black compared to White rates.[84] Second, using data from White Project Implicit respondents aggregated to the county-level, Leitner and colleagues (2016a) found a positive association between Whites' explicit pro-White/anti-Black bias and circulatory disease death rates among Black and White residents, with stronger associations for Black rates.[86]

One study showed that area-level racial prejudice may be protective for Whites. Using Project Implicit data, Orchard & Price (2017) found that the Black-White gap in preterm birth and low birthweight was positively associated with county-level racial prejudice.[90] Most of the racial difference in this association was driven by poorer outcomes among Black birthing persons; however, White birthing persons showed slightly better outcomes in counties with higher levels of explicit anti-Black/pro-White racial prejudice.

Using data from the GSS, Morey et al. (2018) constructed a measure of anti-immigrant "community-level prejudice" within metropolitan statistical areas and non-metropolitan counties.[88] Authors found that community-level anti-immigrant prejudice was not associated with mortality for study respondents overall, nor for White or Black respondents, regardless of whether they were US- or foreign-born. However, US-born "other race" respondents (mostly Hispanics and Asians) living in communities with higher levels of anti-immigrant prejudice had increased mortality rates compared to foreign-born respondents of those ethnicities.

Leitner et al. (2016b) found evidence of race differences that depended on the specification of the study exposure.[87] Authors examined county-level "ingroup" racial bias (i.e., preference for one's own racial group over the other group based on data from Project Implicit) in relation to circulatory disease mortality among Black and White persons. In counties where Black respondents from Project Implicit harbored more implicit ingroup bias (i.e., pro-Black/anti-White), Black residents died from circulatory

disease at a higher rate; whereas in counties where White respondents harbored more explicit ingroup bias (i.e., pro-White/anti-Black), White residents died from circulatory disease at a higher rate.

Using Twitter-characterized measures of racial sentiment, Huang et al. (2020) found race-differential associations that depended on the cardiovascular outcome under consideration.[73] Among non-Hispanic White and Black respondents, Twitter-characterized negative sentiment toward racial/ethnic minoritized groups was most strongly associated with hypertension, myocardial infarction, and CVD. In contrast, among Hispanic respondents, Twitter-characterized negative sentiment toward racial/ethnic minoritized groups was most strongly associated with diabetes, obesity, and stroke.

Findings of Non-Differential Effects by Race/Ethnicity

The remaining four studies that explored effect modification by race/ethnicity showed associations between area-level prejudice and health outcomes that were similar in direction and magnitude (i.e., non-differential) for White as compared with racial/ethnic minoritized groups.[71,74,85,89] Lee et al. (2015) found that area-level prejudice, assessed as a composite measure from the GSS racial attitudes questions, was associated with increased mortality risk among both Black and White GSS respondents.[85] McKetta et al. (2017) found that the state-level popularity of Google searches containing the "n-word" was associated with poor self-rated health among both White and Black Panel Study of Income Dynamics respondents, and no statistically significant interaction with race was observed.[71] Using Twitter data, Nguyen and colleagues found that birthing persons living in states with higher levels of negative,[74] or lower levels of positive,[74,89] racial sentiment had more adverse birth outcomes, with similarly harmful effects observed among racially minoritized and White persons.

Within-group studies

The remaining studies examined outcomes among one racial/ethnic group only, precluding an examination of differential effects.[72,82,83,177] Chae and colleagues found that in DMAs with a greater popularity of Google searches for "n-word," Black residents experienced higher rates of all-cause and cause specific mortality[82] and adverse birth outcomes.[83] Hehman et al. (2018) found that Black residents were killed by police at disproportionate rates relative to their population shares in metropolitan areas where White Project Implicit respondents demonstrated greater implicit and explicit pro-White/anti-Black racial bias.[72] Finally, using Twitter data at the national level, Hswen et al. (2020) found that the percent of negative sentiment tweets mentioning Mexicans and Hispanics was associated with increased daily worry among those identifying as Hispanic during the 20-week period surrounding the 2016 US presidential election.[177]

Theory, conceptualization, and proposed pathways to health

Several studies referenced theories in their introduction or discussion sections, including "structural stigma"[88] and "fundamental causes;"[73] however, there was minimal incorporation of the theories and their constructs in the interpretation of the study findings. Few studies formally defined area-level racial prejudice or how it was understood in relation to existing models of racism.[71,73,88] However, the general framing and stated motivation for the studies provided insight into authors' conceptualization of the exposure.

The conceptualization of area-level racial prejudice fell into two categories: (1) area-level racial prejudice as a proxy for the attitudes and actions of prejudiced individuals, with pathways to health occurring via personally-mediated racial discrimination and its consequences or (2) area-level racial prejudice as a broader measure of prejudiced social contexts, with pathways to health occurring at multiple social-ecologic levels. These two conceptualizations were not mutually-exclusive (i.e., some studies described both) and were not specific to the data source or indicator used (i.e., identical measures of area-level racial prejudice were described as measures of prejudiced individuals or prejudiced contexts depending on the study). Pathways were generally explicated in hypothetical terms either in the introduction to provide social and/or biological plausibility for justifying the research question, or in the discussion to explain findings and make recommendations for future work. These pathways are illustrated in Figure 2.

[FIGURE 2 HERE]

Measuring Prejudiced Individuals

Many authors described racial discrimination as a key determinant of health but noted limitations of traditional survey instruments which ask individuals to recall past experiences of racially motivated discrimination in various settings. These experiences may be underreported on surveys due to the subtlety of contemporary racism, attributional ambiguity, internalized racism, and the influence of coping style on reporting,[74,82,83,86,88,177] which may result in inconsistent or under-estimated effects of racial discrimination on health.[73,74,82,83,89,177]

The use of data from the GSS, Project Implicit, Google Trends, and Twitter was described as a strategy to directly assess the racial prejudice of individuals in a defined geographic area, which may serve as a proxy for exposure to racial discrimination, without relying on study respondents to self-report those experiences.[82,83,86,87] An added benefit of the IAT, Google Trends, and Twitter (as opposed to the explicit measures of racial prejudice from the GSS and Project Implicit) is they do not require prejudiced individuals to openly disclose prejudiced attitudes, which may also be underreported due to self-censorship and social desirability bias.[71,73,74,83,177]

The stated or unstated assumption of these studies was that racial/ethnic minoritized groups living in areas where more individuals harbored racial prejudice would be at increased risk of experiencing racial discrimination. For example, Chae and colleagues (2015, p. 3) noted: "an Internet search-based measure of area racism may serve as a more direct indicator of racial attitudes and the extent of discrimination and prejudice towards Blacks in a geographic area, including those experiences of racially motivated bias that are subtle or not observable, and which are not necessarily reported in survey instruments."

Racial discrimination, in turn, was posited to harm health via biopsychosocial, behavioral, and material mechanisms, as shown in Figure 2. Authors described racial discrimination and intergroup tension as psychosocial stressors that cause increased anger, anxiety, and chronic physiological stress adaption, all of which disrupt key regulatory systems and undermine health.[71,73,74,82,83,85-87,89,90] They also described how this stress can lead to maladaptive coping behaviors, such as poor diet and exercise.[73,74,82,83,85,86,90] Lastly, authors described how racial discrimination in institutional settings (e.g., employment and housing) limits economic opportunities and material resources needed to achieve optimal health,[71,74,82,83,86,89] or harms health directly in the case of healthcare discrimination.[74,83,86,90]

Pathways were described in hypothetical terms but remained largely unexamined, with a few exceptions.[86,90] Orchard & Price (2017) found that birth outcomes were more strongly associated with the racial prejudice in the county of birth compared to the county of residence, hypothesizing that racial bias in county of birth may proxy healthcare discrimination. Leitner et al. (2016a) found that the association between Whites' racial prejudice and circulatory disease inequities was not mediated by Black-White differences in smoking, drinking, and exercise,[86] suggesting non-behavioral factors must be considered.

Measuring Prejudiced Social Contexts

Others argued that the self-reported measures of racial discrimination commonly used in the literature may fail to account for macro-level manifestations of racism,[71-74,84,85,88] and therefore may under-estimate the total effects of racism on health.[73,74,88] In these studies, measuring area-level racial prejudice was motivated by a desire to rise above the level of the individual [75] and capture the broader social context[71-74,84,85,88,89,177] in which prejudiced attitudes and discriminatory behaviors occur.[72,73,89] Authors described area-level racial prejudice as capturing "ambient social attitudes,"[71] "macropsychological characteristics of residents,"[72] or a "temperature of the social environment."[89] Some conceptualized area-level racial prejudice as a direct manifestation of structural[71,88] or cultural[73] racism or stigma, the measurement of which could be used to inform structural interventions.

Authors described multiple pathways through which prejudiced social contexts could undermine health. At the community-level, it was hypothesized that area-level racial prejudice would erode community social capital,[74,84-86,88-90] defined as "the norms of reciprocity, trust and social obligation that are essential for minimizing the risks of poor physical, psychological, or social health."[84] An erosion of community social capital may undermine health through (a) reduced trust, bonding, and norms of mutual reciprocity,[84-86,90] (b) less social and emotional support to buffer stressful life events,[89] and (c) less egalitarian political support and commitment to shared resources, resulting in a lack of investment in policies and programs that could enhance the welfare of community members.[74,84,86] Indeed, Lee et al. (2015) found that community social capital mediated the association between community-level racial prejudice and mortality.[85]

Prejudiced social contexts were also hypothesized to maintain social norms that are permissive of racial discrimination, which harms health by increasing stress and material deprivation.[73,74,86,88] However, authors also described direct effects on health, independent of personally-mediated discrimination. As Nguyen et al. (2020) explained: "the social climate of a place represents a complimentary aspect of racial bias and discrimination that may have its own influence on health, independent of individual-level experiences."[74] Similarly, Morey et al. (2018) described the harms of living in prejudiced communities based not only on cumulative exposure to, but also increased awareness of, racialized attitudes;[88] and Leitner et al. (2016a) emphasized how "hostile community environments" could directly evoke stress and harm health.[86]

Measurement and other methodological considerations

Stated limitations noted across multiple studies, regardless of data source, included: unmeasured confounding;[71,82,83,85,86,90] self-selection into more or less prejudiced environments based on underlying health status;[71,82,83,86-88] the inability to establish temporality;[72,82-85,87,89,90] and unknown mechanisms linking area-level racial prejudice with health outcomes.[71-73,83,84,86-89] In addition, the GSS, Project Implicit, Google Trends, and Twitter each presented a unique set of strengths and limitations, which are summarized below and described in more detail in Supplemental Table 3.

General Social Survey

Primary advantages of the GSS include national representativeness,[85,88] specificity of racial/anti-immigrant prejudice measures, and detailed information on respondent characteristics which could be used for weighting or assessing the prejudices of specific respondents (although none of the reviewed studies did this). Primary disadvantages of the GSS include self-censorship of socially undesirable attitudes,[85] subnational non-representativeness,[84] and relatively less temporal and geographic coverage compared to the other data sources due to not all questions being asked on all survey years.[85]

Project Implicit

Strengths of Project Implicit include the large amount of data and the ability to measure implicit bias without relying on self-report.[86,90] Additional advantages include the option to explore or adjust for multiple social biases (weapons stereotype, age, race, or gender bias, etc.),[72,86,90] examine characteristics of the test-taker for descriptive, weighting, or stratification purposes,[72,86,87,90] and compare implicit versus explicit biases.[72,86,87,90] A major limitation of Project Implicit data is that study respondents are self-selected and may not represent the racial bias in their geographic area of residence.[72,86,87,90] Several studies used post-stratification weights to increase representativeness with respect age and/or sex, reporting similar results with weighted and unweighted data.[86,87,90] Finally, using the conventional method of scoring the Project Implicit tests, it is not possible to discern whether a higher implicit bias score is driven by greater pro-White or higher anti-Black bias.[87]

Google Trends

A major strength of Google Trends is that search data do not rely on self-report and therefore may capture private curiosities and socially undesirable attitudes with more authenticity compared to traditional survey instruments.[71,82,83] Studies using Google data described prior validation work, linking area-level rates of searches for the "n-word" with conceptually related constructs, such as voting[71] and other area-based measures of racial bias.[82] Authors also cited work linking other search queries to real-world outcomes, such as infectious diseases, religiosity, and gun ownership.[71,83] The primary limitation of Google Trends is a low degree of measurement specificity. The demographic makeup of users and context of the search is unknown; searches for the "n-word" may not be motivated by racism.[71,82,83] Despite this potential for measurement error, Google Trends provides a high "signal-to-noise ratio" based on the sheer volume of data available.[82,83]

Twitter

The reach of Twitter was described as a major strength: millions of tweets are sent daily, and a large proportion of users make their profiles public.[73,74,89] Authors also noted that the impersonal nature of Twitter may embolden users to speak more freely about racist attitudes than they may otherwise express during in-person interactions or on a survey.[73,74,89,177] Other benefits of the platform include the range of individuals, groups, and businesses represented,[89] predictive validity across a number of health topics and outcomes,[73,74,89,177] and the ability to capture racial sentiment in real-time.[177] Finally, sentiment analysis allows investigators to determine the overall tone of the Tweets referencing racial groups,[73,74,89,177] offering a major strength over Google Trends data where the context or motivation of the search is entirely unknown.[71,82,83] However, misclassification may occur because sentiment analysis cannot capture sarcasm or humor.[73,74,89] Despite the anonymity of Twitter, self-censorship may still influence what

users are willing to post.[74,89] Finally, demographic characteristics of Twitter users are not reliably discernable.[73]

DISCUSSION

To our knowledge, this is the first systematic literature review of studies measuring area-level racial prejudice and examining associations with health outcomes in the United States. Collectively, the fourteen studies reviewed contribute to our growing understanding of racism as a multilevel determinant of health, while also illuminating the need for conceptual development and continued methodological rigor.

Is area-level racial prejudice harmful to health, and for whom?

This systematic review revealed that area-level racial prejudice, measured using several data sources, was associated with adverse health outcomes—ranging from preterm birth to premature mortality. The magnitude of these associations ranged from subtle to more pronounced; differences in effect sizes could be due to exposure and outcome measurement and specification, estimation procedures, or other study distinctions. More work using standardized measures will be necessary to directly compare findings across studies. It is also important to note that even small effects, when measured at the structural level, can be societally quite meaningful because they are scaled over large populations.

Evidence of differential associations by race/ethnicity was mixed. Some studies found the health harms of area-level racial prejudice to be unique to, or more pronounced among, minoritized racial/ethnic groups compared to Whites,[84,86,88,90] whereas others found associations of similar magnitude among racial/ethnic minoritized groups and Whites.[71,74,85,89] Among studies that explored race/ethnic-specific effects, evidence of racial/ethnic differences did not depend on study outcome or mode of assessment, sample, geographic scale, or other design features. The extent to which these findings are driven by exposure data source is not clear. For example, the three studies that used Project Implicit data and examined differential effects by race found that White people's pro-White/anti-Black racial bias was disproportionately associated with adverse health among Black as compared to White residents.[86,87,90] In contrast, studies using GSS data showed both differential[84,88] and non-differential[85] effects by race/ethnicity. Further research is needed to illuminate the conditions underlying whether and to what extent associations between area-level racial prejudice and health differ between racial/ethnic groups.

Findings of differential effects by race/ethnicity are consistent with an understanding of racism as a system of power and oppression that maintains white supremacy and advances the economic, social, and physical wellbeing of Whites through the subordination of racial/ethnic minoritized groups.[2,5,167] Findings of non-differential effects by race/ethnicity, in contrast, align with an understanding of inequality and white supremacy as harmful to the health of everyone in society.[182,183] Another plausible

explanation for the lack of race-specificity could be unmeasured confounding driving spurious associations of similar magnitude between area-level racial prejudice and health among all racial and ethnic groups. Resolving this mixed evidence is an important avenue for future research and can inform the development of interventions to improve population health and reduce health inequities.

Conceptualization and pathways to health

Our primary conceptual question was: "what are area-level measures of racial prejudice capturing in relation to existing understandings of racism, and what are the pathways to health?" The reviewed studies framed the exposure as either a proxy for the prevalence of prejudiced individuals in a geographic area, or as a macro-level construct, capturing the broader social context over and above the prejudice of individual actors. This latter conceptualization is consistent with Payne et al.'s (2017) "Bias of Crowds" model, which posits that "implicit bias is best understood as a social phenomenon that passes through the minds of individuals, but exists with greater stability in the situations they inhabit."[184] Building consensus around the conceptualization of this exposure is of top priority for grounding future work in this area.

Theories of cultural racism[56,68,76] may advance our understanding of area-level racial prejudice and how it undermines health. Defined as "the instillation of the ideology of inferiority in the values, language, imagery, symbols, and unstated assumptions of the larger society,"[56] cultural racism provides a shared framework through which members of society value different racial/ethnic groups. Cultural racism reflects and reinforces structural racism, including institutional practices and policy regimes, which together shape the distribution of resources by race/ethnicity. At the same time, cultural racism creates the prejudice and stereotypes that undergird racial discrimination, a primary source of psychosocial stress and related health outcomes.[56] Cultural racism may also directly impact health through other forms of racism-related stress, such as vicarious racism (e.g., witnessing or hearing about discrimination against family or friends), vigilance (e.g., bracing oneself in anticipation of experiencing discrimination) and collective experiences (e.g., awareness of discrimination against one's racial/ethnic group, exposure to harmful media representations).[2,76]

Previous research provides insights into the social antecedents and consequences of area-level racial prejudice. For example, evidence suggests that racial prejudice has historical roots in slavery and subsequent structural inequities (e.g., racial residential segregation and Black-White gaps in poverty and social mobility),[185] as well as acute racialized social shocks, such as media coverage of the Black Lives Matter movement[186] and the COVID-19 pandemic.[187] Documented social consequences of area-level racial prejudice include racial inequities in access to healthcare,[86] self-employment,[188] and income.[189] Together, this evidence can be used to inform testable mechanistic hypotheses and deepen the conceptual rigor of future work.

Measurement and other methodological considerations

There were several important limitations that prevent drawing causal conclusions from the evidence reviewed. The majority of studies were cross-sectional, examining exposures and outcomes each at one time point only.[72-74,82-84,86,87,89,90] Given existing theory and evidence to suggest that higher rates of disease can drive up collective biases,[190] reverse-causation may pose a threat to validity.[72] Of note, the few studies that measured health outcomes prospectively provided evidence to support the hypothesis that living in an area with high levels of racial prejudice may be associated with incident morbidity and premature mortality.[71,85,88]

A related concern is residual confounding, or a "mixing of effects"[191] of area-level racial prejudice with unmeasured macro-level factors that may also influence health.[192] While studies controlled for a variety of population-level sociodemographic variables, it is not possible to measure all of the complex structural and historical processes that likely shape collective prejudice, and drive health outcomes. Last, how do we define "area-level" (neighborhood, school district, city, county, MSA, state), and might higher levels of aggregation beyond neighborhoods obscure heterogeneity in prejudice across places? Future work could compare effects of prejudice at different geographic scales or examine cross-level interactions (e.g., county*state) to inform mechanistic hypotheses and multi-level policy interventions.

Across the four data sources used to measure area-level racial prejudice, there were important trade-offs between (1) representativeness, (2) self-censorship, and (3) specificity of measurement (Figure 3). First, the value of aggregating data from multiple individuals to characterize the racial prejudice in a defined geographic area relies on the assumption that those individuals are representative of the populations in the areas where they reside. The Google and Twitter measures have been described as highly representative, given the broad accessibility of internet access today.[74,193] However, according to the Pew Research Center, only about 22% of adults use Twitter,[194] and users tend to be younger, more educated, and more progressive than the general population.[195] In addition, Twitter data are restricted to the 83% of users who make their profiles public.[73,74,196] It is possible that public users and the content of their tweets differ systematically from other users, which poses a threat to validity. Moreover, while almost all (99%) of public tweets can be geotagged at the state-level, only about 3-4% provide latitude and longitude information needed to examine sentiment at smaller geographic scales. The GSS is a nationally-representative survey, but may not be completely representative at smaller geographic scales.[84]

[FIGURE 3 HERE]

Project Implicit data are the least representative because respondents are self-selected, with reasons for participation ranging from class assignments and racial bias trainings or pure curiosity.[197] On average, Project Implicit respondents tend to be younger and comprised of more women than the general population.[198] To mitigate this

concern, some studies implemented post-stratification weights to make county averages more representative based on age and/or sex;[86,87,90] however, this strategy does not address self-selection based on unmeasured factors, including amount of racial bias.[75,86,199] Somewhat reassuringly, validation work by Hehman et al. (2019) demonstrated a high degree of convergent validity of unweighted area-level implicit and explicit racial bias data from Project Implicit with racially charged Google searches and nationally representative racial attitude data from the Pew Research Center.[200]

A second consideration is self-censorship of racial prejudice due to social desirability bias, or individuals' reluctance to report social attitudes they perceive to be unacceptable.[201] The survey-based measures of explicit racial attitudes from Project Implicit and the GSS are highly subject to self-censorship because they are self-reported. A benefit of the implicit association test (IAT) from Project Implicit is that it assesses racial prejudice based on implicit/unconscious cognitive associations rather than self-report, thereby mitigating this concern. A major draw of Google and Twitter is that these measures do not rely on self-reported racist attitudes; they discern this information based on the content of people's search queries and tweets.

Third, specificity of measurement refers to whether the instrument is actually measuring racial prejudice. The IAT from Project Implicit is highly specific: it is a validated measure of racial prejudice and is shown to reliably predict discriminatory behavior across a number of settings;[179,202,203] although there is some measurement ambiguity between pro-White and anti-Black bias.[87] The explicit measure of racial bias from Project Implicit and GSS also offer high specificity because they directly ask respondents' attitudes about race-related topics, but as discussed, are threatened by self-censorship.

While Google and Twitter have high generalizability and low self-censorship compared to the survey-based measures, their validity relies on the assumption that searches and tweets actually reflect racial prejudice. The sentiment analysis used to classify Twitter data may help to evaluate this assumption by enabling the researcher to discern the tone (e.g., positive, negative, or neutral) of the tweets, with the caveat that these algorithms are unable to detect humor, sarcasm, or colloquialisms, potentially resulting in residual misclassification.[73,89] With Google, the content and context of the search query is entirely unknown, and it is plausible that the "n-word" is searched for reasons unrelated to racism. A final threat to measurement specificity, shared by Google and Twitter, is the inability to ascertain the racial identity of the user. Whereas Project Implicit and GSS data allow the researcher to perform stratified aggregation capturing area-level prejudice specifically among Whites (i.e., the racial/ethnic group with the most wealth and decision-making power), Google and Twitter data combine the internet activity of all users. To the extent that racially-charged searching and tweeting means something qualitatively different when performed by a White versus racially minoritized individual, the validity of the exposure will be compromised.

No single data source offered superior representativeness, specificity of measurement, and self-censorship. For example, Project Implicit data are the least representative but the most specific. Google and Twitter minimize self-censorship but in doing so, compromise specificity of measurement. These and other trade-offs are illustrated in Figure 3. An outstanding question is whether aggregated data from various sources measure the same or different exposures. Factor analysis would be one way to evaluate this question, and if an underlying factor is found, a composite score combining data from multiple sources may be useful for more holistically capturing area-level racial prejudice.

Recommendations for future research

This critical review contributes to the growing understanding of racism as a social determinant of health and highlights important directions for future research.

First, there is a pressing need to deepen the conceptual and theoretical rigor of this work. What are these aggregated measures of racial prejudice capturing in relation to the field's current understanding of racism? Might they be conceptualized as a component or bi-product of cultural racism?[56,68,76] Evaluating associations between area-level racial prejudice and other manifestations of cultural racism (e.g., media representations, language and symbols, social norms) would provide useful insight. At the same time, researchers may examine how area-level racial prejudice relates to various dimensions of structural racism (e.g., institutional policies and practices, racial residential segregation), and whether these cultural and structural forces interact to shape social and health inequities.

Second, rooted in theory, future research could incorporate additional data to explore mediation and moderation by community-level demographic, economic, environmental, institutional, and political forces, as well as by individual-level identities, socioeconomic characteristics, and psychosocial, behavioral, and biologic processes. For example, testing whether individuals living in areas with higher levels of racial prejudice report more experiences of racial discrimination and whether such discrimination mediates and/or moderates the effects on health could help to inform hypotheses and targeted interventions.[74,75]

Third, many of the threats to causal inference identified in this review (e.g., unmeasured confounding, temporality issues) could be mitigated by the use of natural experiments and econometric methods. These approaches leverage the timing and/or location of "social shocks" to understand how exogenous changes in the social environment cause changes in health outcomes.[192] For example, research examining the relationship between acute or ongoing racially salient events, changes in area-level racial prejudice, and changes in health outcomes among different racial/ethnic groups can advance our understanding of the causal mechanisms at play, while strengthening the conceptualization of the exposure by identifying its antecedents. In tandem, research linking area-level racial prejudice measures with data from longitudinal cohort

studies with multiple prospective outcome assessments across the lifecourse could identify developmental windows of vulnerability and inform targeted interventions.

Lastly, increased use of big data in public health and social science research will likely lead to innovations in the data sources used to measure area-level racial prejudice (e.g., prejudice expressed on Facebook, Instagram, Craigslist). When evaluating the utility of various data sources, researchers can seek to maximize representativeness and specificity of measurement while minimizing self-censorship. Researchers may also consider validation and factor analysis to strengthen measurement. We caution, however, that rigor and innovation in measurement must be accompanied by equally rigorous theoretical work, as specified in the first recommendation for future research. Additional sources of data will not overcome the need for a clearly defined exposure with solid theoretical grounding.

Strengths and limitations of the systematic review

This preregistered systematic literature review had several notable strengths. Developing comprehensive search strings with consultation from a research librarian and searching across four multidisciplinary databases increased the breadth and depth of our search and the likelihood of identifying eligible articles from across academic disciplines. Two independent investigators performed all screening and data extraction, and met regularly to build consensus, strengthening the quality of our data.

There were also important limitations. Despite a comprehensive search strategy which yielded over 14,000 unique articles, it is possible we did not identify all eligible studies. In addition, publication and reporting bias pose a threat to validity in systematic literature reviews; null or counterintuitive findings may be omitted or unpublished, which would over-estimate the effects of area-level racial prejudice on health.[204] Our working definition of area-level racial prejudice focused on aggregated indicators of cognitive but not behavioral manifestations of racism. This choice was made strategically to manage the conceptual specificity and scope of the study. A future systematic review examining health consequences of aggregated measures of racial discrimination (e.g., home mortgage discrimination,[81] racist policing,[205] or hate crimes)[206] could complement this study.

CONCLUSIONS

The measurement of area-level racial prejudice based on aggregated individual-level data, and estimation of associations with key health outcomes, is a promising area of inquiry. We conducted the first, to our knowledge, systematic review of this growing literature. Evidence to-date suggests that area-level racial prejudice is harmful to the health of racial/ethnic minoritized groups, with some studies also showing harmful effects among Whites. Future studies can deepen the theoretical rigor of this work while advancing innovations in measurement, strengthening causal inference, and exploring social and biologic mechanisms. Findings from this emerging body of

literature can be used to inform public health policymakers and practitioners in developing evidence-based interventions to reduce the harm caused by ambient racial prejudice in society.

TABLES AND FIGURES

Figure 1. Prisma flow diagram

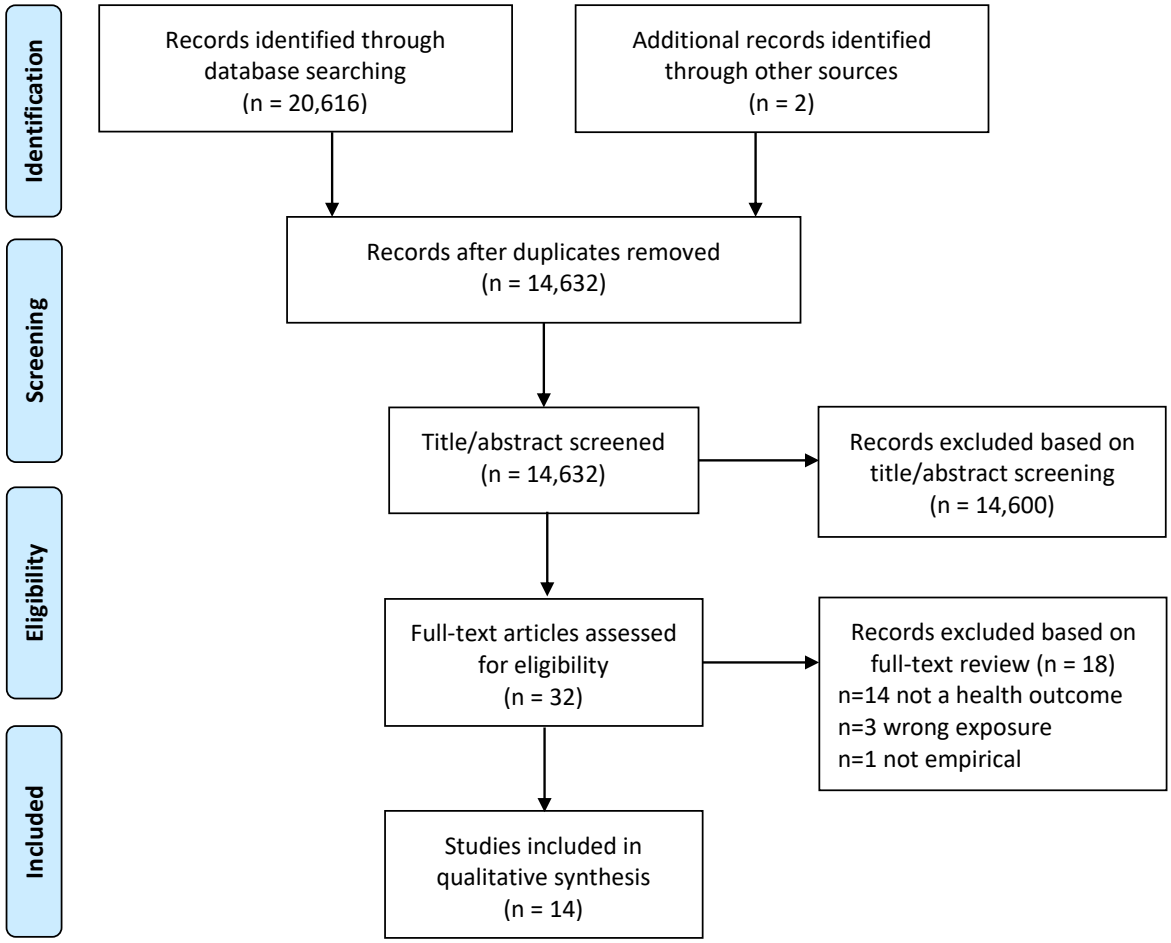


Table 1. Characteristics of included studies (N=14)

Study characteristic	n (%)	References
Follow-up [a]		
Cross-sectional	10 (71.4%)	*Cross-sectional:* Kennedy et al. (1997), Leitner et al. (2016a), Leitner et al. (2016b), Orchard & Price (2017), Hehman (2018), Chae et al. (2015), Chae et al. (2018), Nguyen et al. (2018), Huang et al. (2020) *Serial cross-sectional:* Nguyen et al. (2020)
Prospective	3 (21.4%)	Lee et al. (2015), Morey et al. (2018), McKetta et al. (2017),
Time series	1 (7.1%)	Hswen et al. (2020)
Level of analysis [b]		
Ecologic	6 (42.9%)	Chae et al. (2015), Kennedy et al. (1997), Leitner et al. (2016a), Leitner et al. (2016b), Hehman et al. (2018), Hswen et al. (2020)
Multilevel	7 (50%)	Lee et al. (2015), Morey et al. (2018), McKetta et al. (2017), Orchard & Price (2017), Chae et al. (2018), Nguyen et al. (2018), Nguyen et al. (2020)
Individual	1 (7.1%)	Huang et al. (2020)
Exposure scale		
County	3 (21.4%)	Leitner et al. (2016a), Leitner et al. (2016b), Orchard & Price (2017)
State	5 (35.7%)	Kennedy et al. (1997) McKetta et al. (2017), Nguyen et al. (2018), Nguyen et al. (2020), Huang et al. (2020)
DMA	2 (14.3%)	Chae et al. (2015), Chae et al. (2018)
PSU	2 (14.3%)	Lee et al. (2015), Morey et al. (2018)
CBSA	1 (7.1%)	Hehman et al. (2018)
National	1 (7.1%)	Hswen et al. (2020)
Exposure data source		
GSS	3 (21.4%)	Kennedy et al. (1997), Lee et al. (2015), Morey et al. (2018)
Project Implicit	4 (28.6%)	Leitner et al. (2016a), Leitner et al. (2016b), Orchard & Price (2017), Hehman (2018)
Google Trends	3 (21.4%)	Chae et al. (2015), Chae et al. (2018), McKetta et al. (2017)
Twitter	4 (28.6%)	Nguyen et al. (2018), Nguyen et al. (2020), Huang et al. (2020), Hswen et al. (2020)
Race/ethnicity of exposure population [c]		
White racial prejudice	3 (21.4%)	Leitner et al. (2016a), Leitner et al. (2016b), Hehman et al. (2018)
Black racial prejudice	2 (14.3%)	Leitner et al. (2016b), Hehman et al. (2018)
Not race-specific	4 (28.6%)	Kennedy et al. (1997), Lee et al. (2015), Morey et al. (2018), Orchard & Price (2017)
Not discernable with the data source used	7 (50%)	Chae et al. (2015), Chae et al. (2018), McKetta et al. (2017), Huang et al. (2020), Hswen et al. (2020), Nguyen et al. (2018), Nguyen et al. (2020)
Outcome [c]		
Birth outcomes	4 (28.6%)	Orchard & Price (2017), Chae et al. (2018), Nguyen et al. (2018), Nguyen et al. (2020)
All-cause mortality	4 (28.6%)	Kennedy et al. (1997), Lee et al. (2015), Morey et al. (2018), Chae et al. (2015)

Cause-specific mortality	4 (28.6%)	Leitner et al. (2016a), Leitner et al. (2016b), Hehman (2018), Chae et al. (2015)
CVD and related risk factors	2 (14.3%)	Leitner et al. (2016a), Huang et al. (2020)
Mental health	1 (7.1%)	Hswen et al. (2020)
Self-reported health	1 (7.1%)	McKetta et al. (2017)
Outcome assessment [c]		
Birth records	4 (28.6%)	Orchard & Price (2017), Chae et al. (2018), Nguyen et al. (2018), Nguyen et al. (2020)
Death records	6 (42.9%)	Kennedy et al. (1997), Lee et al. (2015), Morey et al. (2018), Chae et al. (2018), Leitner et al. (2016a), Leitner et al. (2016b)
Self-report	4 (28.6%)	Leitner et al. (2016a), McKetta et al. (2017), Huang et al. (2020), Hswen et al. (2020)
Other	1 (7.1%)	Hehman et al. (2018) (Guardian database)
Race/ethnicity of outcome population [c]		
Black and White	6 (42.9%)	Kennedy et al. (1997), Lee et al. (2015), Leitner et al. (2016a), Leitner et al. (2016b), McKetta et al. (2017), Orchard & Price (2017)
Multi-racial (>2)	4 (28.6%)	Huang et al. (2020), Morey et al. (2018), Nguyen et al. (2018), Nguyen et al. (2020)
Black only	3 (21.4%)	Hehman et al. (2018), Chae et al. (2015), Chae et al. (2018)
Hispanic only	1 (7.1%)	Hswen et al. (2020)
Sample size [d]		
Min	875	Hehman et al. (2018)
Max	31,464,451	Orchard & Price (2017)
Median	3,245,8787.5	
Mean (SD)	5,967,3432.9 (10,530,918.1)	
NA (n (%))	2 (14.3%)	Kennedy et al. (1997), Leitner et al. (2016b)
Mechanisms explored [c]		
Mediation	2 (14.3%)	Lee et al. (2015), Leitner et al. (2016a)
Effect measure modification/interaction [e]	10 (71.4%)	*Compared race-specific rates:* Kennedy et al. (1997) *Stratified subgroups (no test for interaction):* Nguyen et al. (2018), Nguyen et al. (2020) *Multiplicative interaction:* Lee et al. (2015), Morey et al. (2018), Leitner et al. (2016a), Leitner et al. (2016b), Orchard & Price (2017), McKetta et al. (2017), Huang et al. (2020)
NA (within-group study)	4 (28.6%)	Hehman (2018), Chae et al. (2018), Chae et al. (2015), Hswen et al. (2020)

[a] Cross-sectional: exposure and outcome examined at one time-point only, even if the exposure preceded the outcome; Prospective: exposure precedes outcome and there are multiple outcome assessments on each study participant; Time-series analysis: changes in group-level rates (not necessarily the same study participants) over time.

[b] Multilevel: area-level exposure and individual-level outcomes, analyzed using analytic methods that account for clustered data; Individual: exposure and outcome analyzed at individual-level and methods do not account for clustering (i.e., exposure is measured at the area-level but the unit of analysis is the individual); Ecologic: exposure and outcome analyzed at area-level. Individual-level variables may be used in creation of area-level outcome (e.g., age-standardized death rates), but the unit of analysis is the area (e.g., state, county).

[c] Sum is >100% because categories are not mutually exclusive.

[d] Reflects number of individuals included in outcome assessment. NAs are for studies where outcome was expressed as area-level rates and no individual-level sample sizes were provided. Sample sizes and number aggregated in exposure assessments are shown Supplemental Table 1.
[e] Includes any study that examined whether the association between area-level racial prejudice was differential or non-differential by racial/ethnic/immigrant group. This included formal tests of interaction with or without simple slopes or stratified subgroup analysis, stratified subgroup analyses without a formal test of interaction, or examining race-specific rates at the area-level.
DMA = designated market area (geographic area receiving similar media and news); CBSA = core-based statistical area (similar to metropolitan statistical area); PSU = primary sampling unit (comprised of metropolitan statistical areas and non-metropolitan counties).[88]

Figure 2. Proposed pathways linking area-level racial prejudice with health

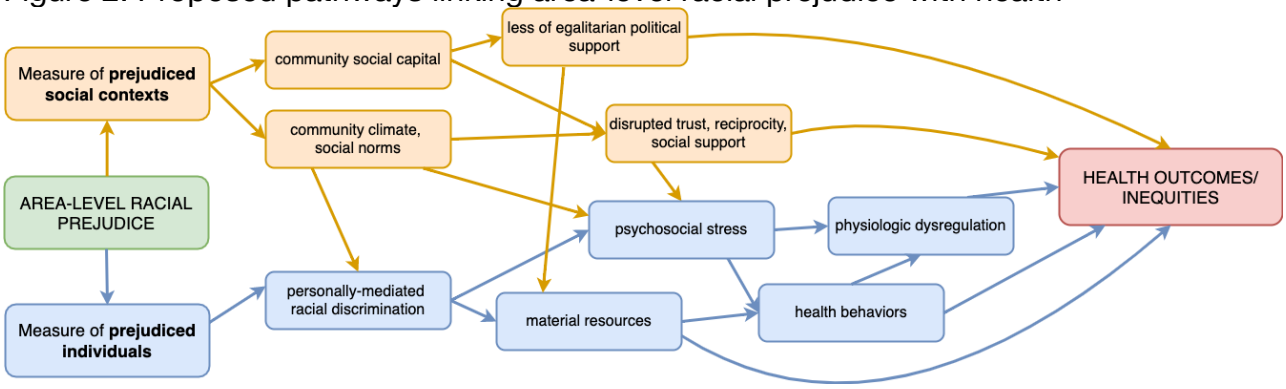
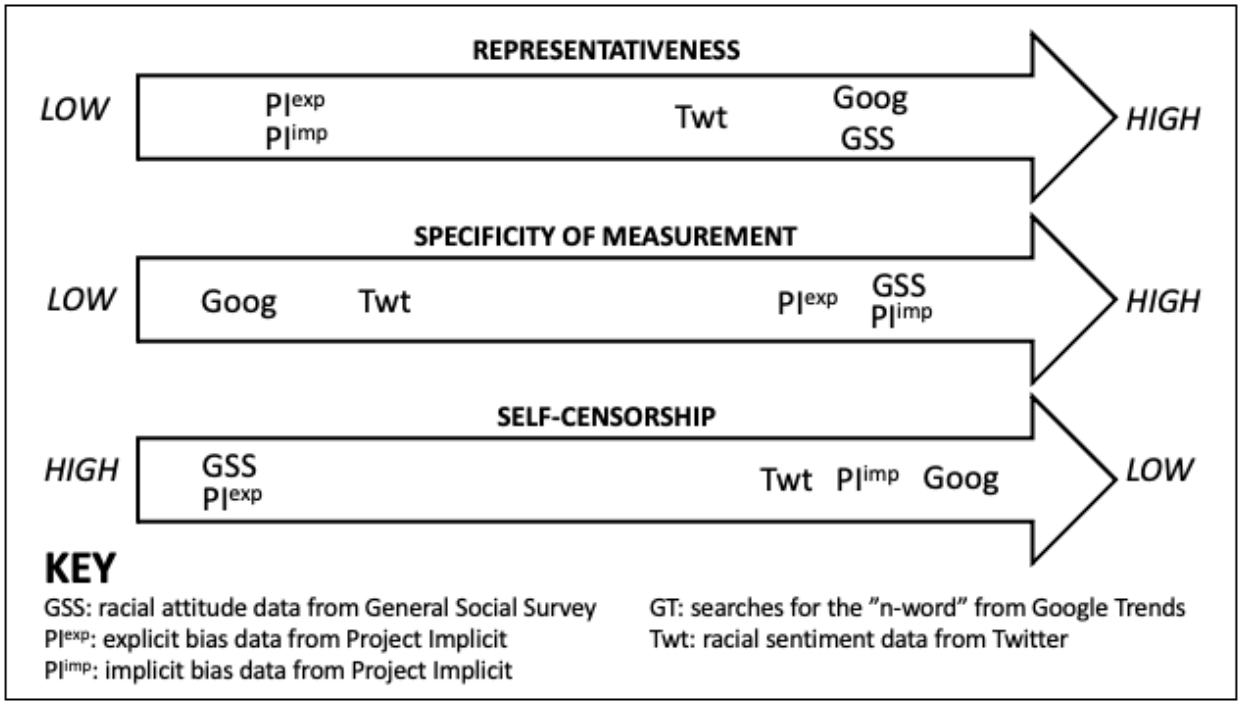


Figure 3. Measurement trade-offs between area-level racial prejudice data sources



CONCLUSION

This dissertation aimed to advance the measurement of racism as a multilevel construct and fundamental cause of health inequities. Drawing on conventional and novel measurement approaches, I examined the health consequences of interpersonal, institutional, and cultural racism. A primary focus of this work was how measurement and methodological decisions can be guided by the conceptualization of the exposure and hypothesized pathways to embodiment. As Cross (2019) poignantly states: "to document racism or study its effects, one must first measure it."[78] This was the charge of my dissertation.

In Chapter 1, we developed a novel approach to coding the Everyday Discrimination Scale with the goal of more accurately capturing the chronicity of racial discrimination experiences, based on the biological pathway to health: chronic psychosocial stress adaptation. Using data from the African American Women's Heart and Health Study, we compared exposure classification and health associations between what we call "chronicity coding" and the two standard coding approaches commonly used in the literature. Consistent with our hypothesis, the chronicity-based coding approach showed the strongest associations with hypertension, which I posited may be due to more accurate exposure assessment. These findings suggest that racial discrimination scale coding decisions can have profound impacts on the classification of risk and documentation of health effects. Mixed evidence of associations between racial discrimination and health outcomes (particularly blood pressure outcomes) across the literature may be, in part, due to differences in scale coding across studies. Future research in this area should choose coding approaches based on the proposed pathway to health and consider adopting a chronicity-based coding approach when the biological mechanism involves chronic stress adaptation.

In Chapter 2, we leveraged data from the Home Mortgage Disclosure Act to measure one pernicious form of institutional racism: racial bias in home mortgage lending. We examined associations with race- and subtype-specific breast cancer incidence among non-Hispanic Black and non-Hispanic White females in California, hypothesizing there would be a positive association with triple-negative breast cancer among Black females. Contrary to our hypothesis, no meaningful association between racial bias in home mortgage lending and breast cancer incidence was observed for either subtype among either racial group (i.e., incidence rate ratios were close to 1). Null findings may have been driven by counteracting harmful and salutary forces in neighborhoods with high rates of racial bias in home mortgage lending, and/or the timing of our exposure assessment in relation to our outcome (i.e., concurrent assessment does not account for latency period). Incorporating residential history data to classify exposure to institutional racism and toxic neighborhood environments during salient etiologic windows across the lifecourse could advance this line of research.

In Chapter 3, we conducted a systematic literature review of all studies examining the health consequences of area-level racial prejudice, which I conceptualize as a measure

of cultural racism. All 14 studies included in the review found a positive association between area-level racial prejudice and adverse health outcomes among racial/ethnic minoritized groups, with four studies also showing a similar association among Whites. In response to the inconsistent conceptualizations of the exposure and hypothesized pathways to health across the studies reviewed, I proposed cultural racism as a unifying construct which may impact health via several distinct mechanisms, offering concrete directions for future research in this area, such as developing and validating a cultural racism index, examining interactions with institutional and structural racism, and empirically testing pathways to health across the lifecourse.

Taken together, these three studies contribute to the growing literature on racism as a fundamental cause of health inequities. A primary theme of my dissertation was to join in the ongoing conversation of how we can best measure racism to understand its effects on health. At its core, my dissertation concludes that measurement decisions should be guided by the conceptual definition and dimension of racism, as well as the hypothesized social or biologic pathway to embodiment.[3] I hope this body of work will encourage the critical reflection and theory-driven measurement required to rigorously document and ultimately eliminate racism's harmful effects on the health of racially minoritized individuals, families, and communities.

REFERENCES

1. Jones CP. Levels of racism: a theoretic framework and a gardener's tale. *American journal of public health.* 2000;90(8):1212.
2. Harrell SP. A multidimensional conceptualization of racism-related stress: implications for the well-being of people of color. *The American journal of orthopsychiatry.* 2000;70(1):42-57.
3. Krieger N. Methods for the scientific study of discrimination and health: an ecosocial approach. *Am J Public Health.* 2012;102(5):936-944.
4. Essed P. *Understanding everyday racism: An interdisciplinary theory.* Vol 2: Sage; 1991.
5. Phelan JC, Link BG. Is racism a fundamental cause of inequalities in health? *Annual Review of Sociology.* 2015;41:311-330.
6. Williams DR, Mohammed SA. Racism and Health I: Pathways and Scientific Evidence. *Am Behav Sci.* 2013;57(8):0002764213487340.
7. Clark R, Anderson NB, Clark VR, Williams DR. Racism as a stressor for African Americans. A biopsychosocial model. *Am Psychol.* 1999;54(10):805-816.
8. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *The Lancet.* 2017;389(10077):1453-1463.
9. Gee GC, Ford CL. Structural racism and health inequities: old issues, new directions. *Du Bois review: social science research on race.* 2011;8(1):115-132.
10. Nazroo JY. The structuring of ethnic inequalities in health: economic position, racial discrimination, and racism. *Am J Public Health.* 2003;93(2):277-284.
11. Link BG, Phelan J. Social conditions as fundamental causes of disease. *Journal of health and social behavior.* 1995:80-94.
12. Link BG, Phelan J. Social conditions as fundamental causes of health inequalities. *Handbook of medical sociology.* 2010;6:3-17.
13. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *J Health Soc Behav.* 2010;51 Suppl:S28-40.
14. Phelan JC, Link BG, Diez-Roux A, Kawachi I, Levin B. "Fundamental causes" of social inequalities in mortality: a test of the theory. *J Health Soc Behav.* 2004;45(3):265-285.
15. Bailey ZD, Krieger N, Agenor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *Lancet (London, England).* 2017;389(10077):1453-1463.
16. Calvin R, Winters K, Wyatt SB, Williams DR, Henderson FC, Walker ER. Racism and cardiovascular disease in African Americans. *The American journal of the medical sciences.* 2003;325(6):315-331.
17. Gee GC, Walsemann KM, Brondolo E. A life course perspective on how racism may be related to health inequities. *Am J Public Health.* 2012;102(5):967-974.
18. Krieger N. Living and Dying at the Crossroads: Racism, Embodiment, and Why Theory Is Essential for a Public Health of Consequence. *Am J Public Health.* 2016;106(5):832-833.

19. Nuru-Jeter A, Dominguez TP, Hammond WP, et al. "It's the skin you're in": African-American women talk about their experiences of racism. an exploratory study to develop measures of racism for birth outcome studies. *Matern Child Health J.* 2009;13(1):29-39.
20. Shavers VL, Shavers BS. Racism and health inequity among Americans. *Journal of the National Medical Association.* 2006;98(3):386.
21. Riddell CA, Morrison KT, Kaufman JS, Harper S. Trends in the contribution of major causes of death to the black-white life expectancy gap by US state. *Health & place.* 2018;52:85-100.
22. Virani SS, Alonso A, Benjamin EJ, et al. Heart disease and stroke statistics—2020 update: a report from the American Heart Association. *Circulation.* 2020;141(9):e139-e596.
23. Virani SS, Alonso A, Aparicio HJ, et al. Heart disease and stroke statistics—2021 update: a report from the American Heart Association. *Circulation.* 2021;143(8):e254-e743.
24. Balfour PC, Rodriguez CJ, Ferdinand KC. The role of hypertension in race-ethnic disparities in cardiovascular disease. *Current cardiovascular risk reports.* 2015;9(4):18.
25. Mottillo S, Filion KB, Genest J, et al. The metabolic syndrome and cardiovascular risk: a systematic review and meta-analysis. *Journal of the American College of Cardiology.* 2010;56(14):1113-1132.
26. Balfour PC, Jr., Rodriguez CJ, Ferdinand KC. The Role of Hypertension in Race-Ethnic Disparities in Cardiovascular Disease. *Current cardiovascular risk reports.* 2015;9(4).
27. Matthews KA, Sowers MF, Derby CA, et al. Ethnic differences in cardiovascular risk factor burden among middle-aged women: Study of Women's Health Across the Nation (SWAN). *American heart journal.* 2005;149(6):1066-1073.
28. Thorpe RJ, Fesahazion RG, Parker L, et al. Accelerated health declines among African Americans in the USA. *Journal of Urban Health.* 2016;93(5):808-819.
29. Sharma S, Malarcher AM, Giles WH, Myers G. Racial, ethnic and socioeconomic disparities in the clustering of cardiovascular disease risk factors. *Ethnicity & disease.* 2004;14(1):43-48.
30. Suglia SF, Clark CJ, Gary-Webb TL. Adolescent obesity, change in weight status, and hypertension: racial/ethnic variations. *Hypertension (Dallas, Tex : 1979).* 2013;61(2):290-295.
31. Clark CJ, Alonso A, Spencer RA, Pencina M, Williams K, Everson-Rose SA. Predicted long-term cardiovascular risk among young adults in the national longitudinal study of adolescent health. *American journal of public health.* 2014;104(12):e108-e115.
32. Shanahan L, Freeman J, Bauldry S. Is very high C-reactive protein in young adults associated with indicators of chronic disease risk? *Psychoneuroendocrinology.* 2014;40:76-85.
33. Hao G, Wang X, Treiber FA, Harshfield G, Kapuku G, Su S. Blood pressure trajectories from childhood to young adulthood associated with cardiovascular

risk: results from the 23-year longitudinal Georgia stress and heart study. *Hypertension (Dallas, Tex : 1979).* 2017;69(3):435-442.
34. Hayward MD, Miles TP, Crimmins EM, Yang Y. The significance of socioeconomic status in explaining the racial gap in chronic health conditions. *American sociological review.* 2000:910-930.
35. Geronimus AT, Bound J, Keene D, Hicken M. Black-white differences in age trajectories of hypertension prevalence among adult women and men, 1999-2002. *Ethnicity & disease.* 2007;17(1):40-49.
36. Geronimus AT, Andersen HF, Bound J. Differences in hypertension prevalence among US black and white women of childbearing age. *Public Health Reports.* 1991;106(4):393.
37. Geronimus AT, Hicken M, Keene D, Bound J. "Weathering" and age patterns of allostatic load scores among blacks and whites in the United States. *Am J Public Health.* 2006;96(5):826-833.
38. DeSantis CE, Siegel RL, Sauer AG, et al. Cancer statistics for African Americans, 2016: progress and opportunities in reducing racial disparities. *CA: a cancer journal for clinicians.* 2016;66(4):290-308.
39. Williams DR, Mohammed SA, Shields AE. Understanding and effectively addressing breast cancer in African American women: Unpacking the social context. *Cancer.* 2016;122(14):2138-2149.
40. Lia Scott LM, Tzy-Mey Kuo, and Dora Il'yasova. Update on triple-negative breast cancer disparities for the United States – a population-based study from the United States Cancer Statistics database, 2010-2014. *Cancer.* 2019(Online first).
41. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA: A Cancer Journal for Clinicians.* 2019.
42. Bureau USC. *American Community Survey 2007-2011 5-year Estimates.* 14 April 2018 2010.
43. Clarke CA, Keegan TH, Yang J, et al. Age-specific incidence of breast cancer subtypes: understanding the black–white crossover. *Journal of the National Cancer Institute.* 2012;104(14):1094-1101.
44. Polyak K. Heterogeneity in breast cancer. *The Journal of clinical investigation.* 2011;121(10):3786-3788.
45. DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast cancer statistics, 2017, racial disparity in mortality by state. *CA: a cancer journal for clinicians.* 2017;67(6):439-448.
46. DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, Jemal A. Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA Cancer J Clin.* 2016;66(1):31-42.
47. Martin JA, Hamilton BE, Osterman M. Births in the United States, 2020. *NCHS data brief.* 2021(418):1-8.
48. Cowie CC, Casagrande SS, Geiss LS. Prevalence and incidence of type 2 diabetes and prediabetes. 2021.

49. Wang MC, Shah NS, Carnethon MR, O'Brien MJ, Khan SS. Age at Diagnosis of Diabetes by Race and Ethnicity in the United States From 2011 to 2018. *JAMA internal medicine.* 2021.
50. Neighbors HW, Caldwell C, Williams DR, et al. Race, ethnicity, and the use of services for mental disorders: results from the National Survey of American Life. *Arch Gen Psychiatry.* 2007;64(4):485-494.
51. Farkas K, dP Duarte C, Ahern J. Injuries to children and adolescents by law enforcement: an analysis of California emergency department visits and hospitalizations, 2005-2017. *JAMA pediatrics.* 2021.
52. Edwards F, Lee H, Esposito M. Risk of being killed by police use of force in the United States by age, race–ethnicity, and sex. *Proceedings of the National Academy of Sciences.* 2019;116(34):16793-16798.
53. Gee GC. A multilevel analysis of the relationship between institutional and individual racial discrimination and health status. *American journal of public health.* 2008;98(Supplement_1):S48-S56.
54. Reskin B. The race discrimination system. *Annual Review of Sociology.* 2012;38:17-35.
55. Powell JA. Structural racism: Building upon the insights of John Calmore. *NCL Rev.* 2007;86:791.
56. Williams DR, Lawrence JA, Davis BA. Racism and health: evidence and needed research. *Annual review of public health.* 2019;40:105-125.
57. Gee G, Hicken, M. Commentary – Structural Racism: The Rules and Relations of Inequity. *Ethnicity & Disease.* 2021;Volume 31:1-8.
58. Kessler RC, Mickelson KD, Williams DR. The Prevalence, Distribution, and Mental Health Correlates of Perceived Discrimination in the United States. *Journal of Health and Social Behavior.* 1999;40(3):208.
59. Williams DR, Neighbors H. Racism, discrimination and hypertension: evidence and needed research. *Ethn Dis.* 2001;11(4):800-816.
60. Folkman La. Stress, Appraisal, and Coping. *New York.* 1984;1.
61. McEwen BS. Stress, adaptation, and disease. Allostasis and allostatic load. *Ann N Y Acad Sci.* 1998;840(1):33-44.
62. Krieger N, Smith K, Naishadham D, Hartman C, Barbeau EM. Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. *Soc Sci Med.* 2005;61(7):1576-1596.
63. Duarte C.dP. B, M., Moses, C., Kajeepeta, S., Prins, S., Scott, J., Mujahid, M. A transdisciplinary systematic literature review of school discipline and health studies in US schools: Summary of findings and future directions for research and practice. 2021.
64. Duarte CdP, Salas-Hernández L, Griffin JS. Policy Determinants of Inequitable Exposure to the Criminal Legal System and Their Health Consequences Among Young People. *American Journal of Public Health.* 2020;110(S1):S43-S49.
65. Alexander M. The new jim crow. *Ohio St J Crim L.* 2011;9:7.
66. Bailey ZD, Feldman JM, Bassett MT. How structural racism works—racist policies as a root cause of US racial health inequities. In*.* Vol 384: Mass Medical Soc; 2021:768-773.

67. Hicken MT, Miles L, Haile S, Esposito M. Linking history to contemporary state-sanctioned slow violence through cultural and structural racism. *The Annals of the American Academy of Political and Social Science.* 2021;694(1):48-58.
68. Cogburn CD. Culture, race, and health: implications for racial inequities and population health. *The Milbank Quarterly.* 2019;97(3):736-761.
69. Nazroo JY, Bhui KS, Rhodes J. Where next for understanding race/ethnic inequalities in severe mental illness? Structural, interpersonal and institutional racism. *Sociology of Health & Illness.* 2020;42(2):262-276.
70. Geronimus AT, James SA, Destin M, et al. Jedi public health: Co-creating an identity-safe culture to promote health equity. *SSM-population health.* 2016;2:105-116.
71. McKetta S, Hatzenbuehler ML, Pratt C, Bates L, Link BG, Keyes KM. Does social selection explain the association between state-level racial animus and racial disparities in self-rated health in the United States? *Annals of epidemiology.* 2017;27(8):485-492. e486.
72. Hehman E, Flake JK, Calanchini J. Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Social Psychological and Personality Science.* 2018:1948550617711229.
73. Huang D, Huang Y, Adams N, Nguyen TT, Nguyen QC. Twitter-Characterized Sentiment Towards Racial/Ethnic Minorities and Cardiovascular Disease (CVD) Outcomes. *Journal of Racial and Ethnic Health Disparities.* 2020:1-13.
74. Nguyen TT, Adams N, Huang D, Glymour MM, Allen AM, Nguyen QC. The Association Between State-Level Racial Attitudes Assessed From Twitter Data and Adverse Birth Outcomes: Observational Study. *JMIR Public Health and Surveillance.* 2020;6(3):e17103.
75. Blair IV, Brondolo E. Moving beyond the individual: Community-level prejudice and health. *Social Science & Medicine.* 2017(183):169-172.
76. Hicken M, Stanton, A., Lee, H. The Burden of Cultural Racism: Vigilance and Racial Health Inequalities. In: Ford C, Griffith, D., Bruce, M., Gilbert, K., ed. *Racism: Science & Tools for the Public Health Professional.* Washington, DC: American Public Health Association; 2019:583.
77. Savitz DA. *Interpreting Epidemiologic Evidence: Strategy for Study Design and Analysis.* Oxford, UK: Oxford University Press; 2003.
78. Cross RI. Appendix B. Selected Measures of Racism. In: Ford CL, Griffith DM, Bruce MA, Gilbert KL, eds. *Racism: Science & Tools for the Public Health Professional.* Washington, DC: American Public Health Association; 2019.
79. Williams DR, Yan Y, Jackson JS, Anderson NB. Racial Differences in Physical and Mental Health: Socio-economic Status, Stress and Discrimination. *J Health Psychol.* 1997;2(3):335-351.
80. Krieger N, Smith K, Naishadham D, Hartman C, Barbeau EM. Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. *Social science & medicine.* 2005;61(7):1576-1596.
81. Beyer KM, Zhou Y, Matthews K, Bemanian A, Laud PW, Nattinger AB. New spatially continuous indices of redlining and racial bias in mortgage lending:

links to survival after breast cancer diagnosis and implications for health disparities research. *Health Place.* 2016;40:34-43.
82. Chae DH, Clouston S, Hatzenbuehler ML, et al. Association between an internet-based measure of area racism and Black mortality. *PloS one.* 2015;10(4):e0122963.
83. Chae DH, Clouston S, Martz CD, et al. Area racism and birth outcomes among Blacks in the United States. *Social Science & Medicine.* 2018;199:49-55.
84. Kennedy BP, Kawachi I, Lochner K, Jones C, Prothrow-Stith D. (Dis) respect and black mortality. *Ethnicity & disease.* 1997;7(3):207-214.
85. Lee Y, Muennig P, Kawachi I, Hatzenbuehler ML. Effects of racial prejudice on the health of communities: a multilevel survival analysis. *American journal of public health.* 2015;105(11):2349-2355.
86. Leitner JB, Hehman E, Ayduk O, Mendoza-Denton R. Blacks' death rate due to circulatory diseases is positively related to whites' explicit racial bias: A nationwide investigation using project implicit. *Psychological science.* 2016;27(10):1299-1311.
87. Leitner JB, Hehman E, Ayduk O, Mendoza-Denton R. Racial bias is associated with ingroup death rate for Blacks and Whites: Insights from Project Implicit. *Social Science & Medicine.* 2016;170:220-227.
88. Morey BN, Gee GC, Muennig P, Hatzenbuehler ML. Community-level prejudice and mortality among immigrant groups. *Social Science & Medicine.* 2018;199:56-66.
89. Nguyen TT, Meng H-W, Sandeep S, et al. Twitter-derived measures of sentiment towards minorities (2015–2016) and associations with low birth weight and preterm birth in the United States. *Computers in Human Behavior.* 2018.
90. Orchard J, Price J. County-level racial prejudice and the black-white gap in infant health outcomes. *Social Science & Medicine.* 2017;181:191-198.
91. Krieger N. Theories for social epidemiology in the 21st century: an ecosocial perspective. *Int J Epidemiol.* 2001;30(4):668-677.
92. Linnenbringer E, Gehlert S, Geronimus AT. Black-White Disparities in Breast Cancer Subtype: The Intersection of Socially Patterned Stress and Genetic Expression. *AIMS public health.* 2017;4(5):526.
93. Writing Group M, Mozaffarian D, Benjamin EJ, et al. Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association. *Circulation.* 2016;133(4):e38-360.
94. Statistics NCfH. Health, United States, 2015: with special feature on racial and ethnic health disparities. 2016.
95. Slaughter-Acey JC, Sealy-Jefferson S, Helmkamp L, et al. Racism in the form of micro aggressions and the risk of preterm birth among black women. *Annals of epidemiology.* 2016;26(1):7-13. e11.
96. Paradies Y. A systematic review of empirical research on self-reported racism and health. *Int J Epidemiol.* 2006;35(4):888-901.
97. Harrell JP, Hall S, Taliaferro J. Physiological responses to racism and discrimination: an assessment of the evidence. *Am J Public Health.* 2003;93(2):243-248.

98. Williams DR, Neighbors HW, Jackson JS. Racial/ethnic discrimination and health: findings from community studies. *Am J Public Health.* 2008;98(9 Suppl):S29-37.
99. Hunte HE, King K, Hicken M, Lee H, Lewis TT. Interpersonal discrimination and depressive symptomatology: examination of several personality-related characteristics as potential confounders in a racial/ethnic heterogeneous adult sample. *BMC Public Health.* 2013;13:1084.
100. Roberts CB, Vines AI, Kaufman JS, James SA. Cross-sectional association between perceived discrimination and hypertension in African-American men and women: the Pitt County Study. *Am J Epidemiol.* 2008;167(5):624-632.
101. Cozier Y, Palmer JR, Horton NJ, Fredman L, Wise LA, Rosenberg L. Racial discrimination and the incidence of hypertension in US black women. *Ann Epidemiol.* 2006;16(9):681-687.
102. Schulz AJ, Gravlee CC, Williams DR, Israel BA, Mentz G, Rowe Z. Discrimination, symptoms of depression, and self-rated health among african american women in detroit: results from a longitudinal analysis. *Am J Public Health.* 2006;96(7):1265-1270.
103. Hudson DL, Neighbors HW, Geronimus AT, Jackson JS. Racial Discrimination, John Henryism, and Depression Among African Americans. *Journal of Black Psychology.* 2016;42(3):221-243.
104. Dolezsar CM, McGrath JJ, Herzig AJ, Miller SB. Perceived racial discrimination and hypertension: a comprehensive systematic review. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association.* 2014;33(1):20-34.
105. Brondolo E, Love EE, Pencille M, Schoenthaler A, Ogedegbe G. Racism and hypertension: a review of the empirical evidence and implications for clinical practice. *Am J Hypertens.* 2011;24(5):518-529.
106. Lewis TT, Barnes LL, Bienias JL, Lackland DT, Evans DA, Mendes de Leon CF. Perceived discrimination and blood pressure in older African American and white adults. *J Gerontol A Biol Sci Med Sci.* 2009;64(9):1002-1008.
107. Chae DH, Nuru-Jeter AM, Adler NE. Implicit racial bias as a moderator of the association between racial discrimination and hypertension: a study of Midlife African American men. *Psychosom Med.* 2012;74(9):961-964.
108. Sims M, Diez-Roux AV, Dudley A, et al. Perceived discrimination and hypertension among African Americans in the Jackson Heart Study. *American Journal of Public Health.* 2012;102(S2):S258-S265.
109. Brown C, Matthews KA, Bromberger JT, Chang Y. The relation between perceived unfair treatment and blood pressure in a racially/ethnically diverse sample of women. *Am J Epidemiol.* 2006;164(3):257-262.
110. Moody DLB, Chang YF, Pantesco EJ, et al. Everyday Discrimination Prospectively Predicts Blood Pressure Across 10 Years in Racially/Ethnically Diverse Midlife Women: Study of Women's Health Across the Nation. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine.* 2018.

111. Albert MA, Ravenell J, Glynn RJ, Khera A, Halevy N, de Lemos JA. Cardiovascular risk indicators and perceived race/ethnic discrimination in the Dallas Heart Study. *Am Heart J.* 2008;156(6):1103-1109.
112. Allen AM, Thomas MD, Michaels EK, et al. Racial discrimination, educational attainment, and biological dysregulation among midlife African American women. *Psychoneuroendocrinology.* 2019;99:225-235.
113. Rubin DB. *Multiple imputation for nonresponse in surveys.* Vol 81. New York, NY: John Wiley & Sons; 2004.
114. StataCorp L. Stata multiple-imputation reference manual. 1985.
115. Pickering TG, Hall JE, Appel LJ, et al. Recommendations for blood pressure measurement in humans and experimental animals: part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. *Circulation.* 2005;111(5):697-716.
116. James PA, Oparil S, Carter BL, et al. 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8). *JAMA.* 2014;311(5):507-520.
117. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Applied psychological measurement.* 1977;1(3):385-401.
118. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *Am J Prev Med.* 1994;10(2):77-84.
119. Kohout FJ, Berkman LF, Evans DA, Cornoni-Huntley J. Two shorter forms of the CES-D (Center for Epidemiological Studies Depression) depression symptoms index. *Journal of aging and health.* 1993;5(2):179-193.
120. John OP, Srivastava S. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research.* 1999;2(1999):102-138.
121. Huebner DM, Nemeroff CJ, Davis MC. Do hostility and neuroticism confound associations between perceived discrimination and depressive symptoms? *Journal of Social and Clinical Psychology.* 2005;24(5):723-740.
122. Goodman LA, Kruskal WH. Measures of association for cross classifications. *Journal of the American statistical association.* 1954;49(268):732-764.
123. Cohen J. A coefficient of agreement for nominal scales. *Educational and psychological measurement.* 1960;20(1):37-46.
124. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC medical research methodology.* 2003;3(1):21.
125. StataCorp L. Stata data analysis and statistical Software. *Special Edition Release.* 2007;10:733.
126. Krieger N, Sidney S. Racial discrimination and blood pressure: the CARDIA Study of young black and white adults. *Am J Public Health.* 1996;86(10):1370-1378.
127. Sapolsky RM. *Why zebras don't get ulcers.* WH Freeman New York; 1994.

128. LaVeist TA, Sellers R, Neighbors HW. Perceived racism and self and system blame attribution: consequences for longevity. *Ethn Dis.* 2001;11(4):711-721.
129. Lewis TT, Yang FM, Jacobs EA, Fitchett G. Racial/ethnic differences in responses to the everyday discrimination scale: a differential item functioning analysis. *Am J Epidemiol.* 2012;175(5):391-401.
130. Clark R, Coleman AP, Novak JD. Brief report: Initial psychometric properties of the everyday discrimination scale in black adolescents. *Journal of adolescence.* 2004;27(3):363-368.
131. Bureau U. American community survey. *Raleigh-Durham-Chapel Hill, NC Combined Statistical Area.* 2013.
132. Alegria M, Jackson JS, Kessler RC, Takeuchi D. Collaborative Psychiatric Epidemiology Surveys (CPES), 2001-2003 [United States]. In: Inter-university Consortium for Political and Social Research [distributor]; 2016.
133. Williams DR, Gonzalez HM, Neighbors H, et al. Prevalence and distribution of major depressive disorder in African Americans, Caribbean blacks, and non-Hispanic whites: results from the National Survey of American Life. *Arch Gen Psychiatry.* 2007;64(3):305-315.
134. Assari S. High Income Protects Whites but Not African Americans against Risk of Depression. *Healthcare (Basel, Switzerland).* 2018;6(2).
135. Nuru-Jeter AM, Michaels EK, Thomas MD, Reeves AN, Thorpe RJ, Jr., LaVeist TA. Relative Roles of Race Versus Socioeconomic Position in Studies of Health Inequalities: A Matter of Interpretation. *Annu Rev Public Health.* 2018;39(0):169-188.
136. Warner ET, Gomez SL. Impact of neighborhood racial composition and metropolitan residential segregation on disparities in breast cancer stage at diagnosis and survival between black and white women in California. *Journal of community health.* 2010;35(4):398-408.
137. Beyer KM, Young S, Bemanian A. Persistent Racial Disparities in Breast Cancer Mortality Between Black and White Women: What is the Role for Structural Racism? In: *Geospatial Approaches to Energy Balance and Breast Cancer.* Springer; 2019:361-378.
138. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. *Public Health Rep.* 2001;116(5):404-416.
139. Massey DS, Denton NA. *American apartheid: Segregation and the making of the underclass.* Harvard University Press; 1993.
140. Geronimus AT. To mitigate, resist, or undo: addressing structural influences on the health of urban populations. *American journal of public health.* 2000;90(6):867.
141. Russell EF, Kramer MR, Cooper HL, Gabram-Mendola S, Senior-Crosby D, Arriola KRJ. Metropolitan area racial residential segregation, neighborhood racial composition, and breast cancer mortality. *Cancer Causes & Control.* 2012;23(9):1519-1527.
142. Hossain F, Danos D, Prakash O, et al. Neighborhood Social Determinants of Triple Negative Breast Cancer. *Frontiers in public health.* 2019;7.

143. Krieger N, Feldman JM, Kim R, Waterman PD. Cancer incidence and multilevel measures of residential economic and racial segregation for cancer registries. *JNCI Cancer Spectrum.* 2018;2(1):pky009.
144. Zhou Y, Bemanian A, Beyer KM. Housing discrimination, residential racial segregation, and colorectal cancer survival in southeastern Wisconsin. In: AACR; 2017.
145. Mendez DD, Hogan VK, Culhane JF. Institutional racism, neighborhood factors, stress, and preterm birth. *Ethnicity & health.* 2014;19(5):479-499.
146. Collin LJ, Gaglioti AH, Beyer KM, et al. Neighborhood-level redlining and lending bias are associated with breast cancer mortality in a large and diverse metropolitan area. *Cancer Epidemiology and Prevention Biomarkers.* 2020;30(1):53-60.
147. McCoy P. The home mortgage disclosure act: A synopsis and recent legislative history. *Journal of Real Estate Research.* 2007;29(4):381-397.
148. Krieger N. History, biology, and health inequities: emergent embodied phenotypes and the illustrative case of the breast cancer estrogen receptor. *American journal of public health.* 2013;103(1):22-27.
149. Krieger N, Jahn JL, Waterman PD. Jim Crow and estrogen-receptor-negative breast cancer: US-born black and white non-Hispanic women, 1992-2012. *Cancer Causes Control.* 2017;28(1):49-59.
150. Barnard ME, Boeke CE, Tamimi RM. Established breast cancer risk factors and risk of intrinsic tumor subtypes. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer.* 2015;1856(1):73-85.
151. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *Jama.* 2006;295(21):2492-2502.
152. Gaudet MM, Gierach GL, Carter BD, et al. Pooled analysis of nine cohorts reveals breast cancer risk factors by tumor molecular subtype. *Cancer research.* 2018;78(20):6011-6021.
153. Holm J, Eriksson L, Ploner A, et al. Assessment of breast cancer risk factors reveals subtype heterogeneity. *Cancer research.* 2017;77(13):3708-3717.
154. Wheeler SB, Reeder-Hayes KE, Carey LA. Disparities in Breast Cancer Treatment and Outcomes: Biological, Social, and Health System Determinants and Opportunities for Research. *The Oncologist.* 2013;18(9):986-993.
155. Ellingjord-Dale M, Vos L, Tretli S, Hofvind S, dos-Santos-Silva I, Ursin G. Parity, hormones and breast cancer subtypes-results from a large nested case-control study in a national screening program. *Breast cancer research.* 2017;19(1):10.
156. Yang XR, Chang-Claude J, Goode EL, et al. Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. *Journal of the National Cancer Institute.* 2010;103(3):250-263.
157. Bureau UC. American community survey. In: US Department of Commerce, Economics and Statistics Administration, US Census Bureau Washington, DC; 2010.

158. Yost K, Perkins C, Cohen R, Morris C, Wright W. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer Causes & Control.* 2001;12(8):703-711.
159. Yang J, Schupp C, Harrati A, Clarke C, Keegan T, Gomez S. Developing an area-based socioeconomic measure from American Community Survey data. *Fremont, CA: Cancer Prevention Institute of California.* 2014.
160. Fenton SE, Birnbaum LS. Timing of environmental exposures as a critical element in breast cancer risk. *The Journal of Clinical Endocrinology & Metabolism.* 2015;100(9):3245-3250.
161. Gomez SL, Shariff-Marco S, DeRouen M, et al. The impact of neighborhood social and built environment factors across the cancer continuum: current research, methodological considerations, and future directions. *Cancer.* 2015;121(14):2314-2330.
162. Deardorff J, Fyfe M, Ekwaru JP, Kushi LH, Greenspan LC, Yen IH. Does neighborhood environment influence girls' pubertal onset? findings from a cohort study. *BMC pediatrics.* 2012;12(1):1-9.
163. Hurley SE, Reynolds P, Goldberg DE, et al. Residential mobility in the California Teachers Study: implications for geographic differences in disease rates. *Social science & medicine.* 2005;60(7):1547-1555.
164. Krieger N, Singh N, Waterman PD. Metrics for monitoring cancer inequities: residential segregation, the Index of Concentration at the Extremes (ICE), and breast cancer estrogen receptor status (USA, 1992–2012). *Cancer Causes & Control.* 2016;27(9):1139-1151.
165. Linnenbringer E, Geronimus AT, Davis KL, Bound J, Ellis L, Gomez SL. Associations between breast cancer subtype and neighborhood socioeconomic and racial composition among Black and White women. *Breast Cancer Research and Treatment.* 2020:1-11.
166. Martínez ME, Cruz GI, Brewster AM, Bondy ML, Thompson PA. What can we learn about disease etiology from case-case analyses? Lessons from breast cancer. *Cancer Epidemiology and Prevention Biomarkers.* 2010;19(11):2710-2714.
167. Feagin J. *Systemic racism: A theory of oppression.* Routledge; 2013.
168. Chetty R, Hendren N. The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *The Quarterly Journal of Economics.* 2018;133(3):1107-1162.
169. Mooney SJ, Westreich DJ, El-Sayed AM. Epidemiology in the era of big data. *Epidemiology (Cambridge, Mass).* 2015;26(3):390.
170. Stephens-Davidowitz S, Pabon A. *Everybody lies: Big data, new data, and what the internet can tell us about who we really are.* HarperCollins New York; 2017.
171. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS med.* 2009;6(7):e1000097.
172. *Covidence systematic review software* [computer program]. Melbourne, Australia2016.

173. Novak NL, Geronimus AT, Martinez-Cardoso AM. Change in birth outcomes among infants born to Latina mothers after a major immigration raid. *International journal of epidemiology.* 2017;46(3):839-849.
174. Gee GC, Ro A, Shariff-Marco S, Chae D. Racial discrimination and health among Asian Americans: evidence, assessment, and directions for future research. *Epidemiologic reviews.* 2009;31(1):130-151.
175. Viruell-Fuentes EA, Miranda PY, Abdulrahim S. More than culture: structural racism, intersectionality theory, and immigrant health. *Social science & medicine.* 2012;75(12):2099-2106.
176. *MAXQDA 2020* [computer program]. Berlin, Germany2019.
177. Hswen Y. Online negative sentiment towards Mexicans and Hispanics and impact on mental well-being: A time-series analysis of social media data during the 2016 United States presidential election. *Heliyon.* 2020;6(9).
178. Davis JA, Schwartzman K. *General Social Survey: March 1975.* Vol 4: Inter-University Consortium for Political Research; 1973.
179. Nosek B, Banaji M, Greenwald A. Project implicit. *Project Implicit.* 2010.
180. Google. Google Trends. https://trends.google.com/trends/?geo=US. Published 2020. Accessed.
181. Jewell NP. *Statistics for epidemiology.* Chapman and Hall/CRC; 2003.
182. Wilkinson RG. *Mind the gap: hierarchies, health and human evolution.* Yale University Press; 2001.
183. Malat J, Mayorga-Gallo S, Williams DR. The effects of whiteness on the health of whites in the USA. *Social Science & Medicine.* 2018;199:148-156.
184. Payne BK, Vuletich HA, Lundberg KB. The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry.* 2017;28(4):233-248.
185. Payne BK, Vuletich HA, Brown-Iannuzzi JL. Historical roots of implicit bias in slavery. *Proceedings of the National Academy of Sciences.* 2019;116(24):11693-11698.
186. Sawyer J, Gampa A. Implicit and Explicit Racial Attitudes Changed During Black Lives Matter. *Personality and Social Psychology Bulletin.* 2018:0146167218757454.
187. Darling-Hammond S, Michaels EK, Allen AM, et al. After "The China Virus" Went Viral: Racially Charged Coronavirus Coverage and Trends in Bias Against Asian Americans. *Health Education & Behavior.* 2020.
188. Kopkin N. Does racial prejudice affect black entrepreneurship?: evidence exploiting spatial differences in prejudicial attitudes. *Applied Economics.* 2017;49(31):3045-3066.
189. Connor P, Sarafidis V, Zyphur MJ, Keltner D, Chen S. Income inequality and White-on-Black racial bias in the United States: Evidence from project implicit and Google Trends. *Psychological science.* 2019;30(2):205-222.
190. O'Shea BA, Watson DG, Brown GD, Fincher CL. Infectious disease prevalence, not race exposure, predicts both implicit and explicit racial prejudice across the United States. *Social Psychological and Personality Science.* 2020;11(3):345-355.

191. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology.* Lippincott Williams & Wilkins; 2008.
192. Glymour MM. Natural experiments and instrumental variable analyses in social epidemiology. *Methods in social epidemiology.* 2006;1:429.
193. Stephens-Davidowitz S. The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics.* 2014;118:26-40.
194. Perrin A, Anderson M. Share of US adults using social media, including Facebook, is mostly unchanged since 2018. *Pew Research Center.* 2019;10.
195. Wojcik S, Hughes A. Sizing up Twitter users. *Washington, DC: Pew Research Center.* 2019.
196. Remy E. How public and private Twitter users in the U.S. compare — and why it might matter for your research. *Decoded.* 2019. https://medium.com/pew-research-center-decoded/how-public-and-private-twitter-users-in-the-u-s-d536ce2a41b3.
197. Sakong J. Identifying taste-based discrimination: Effect of black electoral victories on racial prejudice and economic gaps. *Federal Reserve Bank of Chicago.* 2021;Working Paper, No. 2021-07.
198. Xu K, Nosek, B. and Greenwald, A.G. Psychology data from the Race Implicit Association Test on the Project Implicit Demo website. In. *Journal of Open Psychology Data, 2(1), p.e3.*2014.
199. Hoover J, Dehghani M. The big, the bad, and the ugly: Geographic estimation with flawed psychological data. *Psychological Methods.* 2020;25(4):412-429.
200. Hehman E, Calanchini J, Flake JK, Leitner JB. Establishing construct validity evidence for regional measures of explicit and implicit racial bias. *Journal of Experimental Psychology: General.* 2019;148(6):1022-1040.
201. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity.* 2013;47(4):2025-2047.
202. Greenwald AG, Poehlman TA, Uhlmann EL, Banaji MR. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of personality and social psychology.* 2009;97(1):17.
203. Jost JT, Rudman LA, Blair IV, et al. The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore. *Research in organizational behavior.* 2009;29:39-69.
204. Dwan K, Altman DG, Arnaiz JA, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PloS one.* 2008;3(8):e3081.
205. Sewell AA, Jefferson KA. Collateral damage: the health effects of invasive police encounters in New York City. *Journal of Urban Health.* 2016;93(1):42-67.
206. Duncan DT, Hatzenbuehler ML. Lesbian, gay, bisexual, and transgender hate crimes and suicidality among a population-based sample of sexual-minority adolescents in Boston. *American journal of public health.* 2014;104(2):272-278.
207. (CDC) CfDCaP. Behavioral Risk Factor Surveillance System Survey. In: Services USDoHaH, ed. Atlanta, GA.

208. Center SR. Panel Study of Income Dynamics. In: Institute for Social Research UoM, ed. Ann Arbor, MI.
209. Greenwald AG, Nosek BA, Banaji MR. Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of personality and social psychology.* 2003;85(2):197.

SUPPLEMENTAL MATERIAL

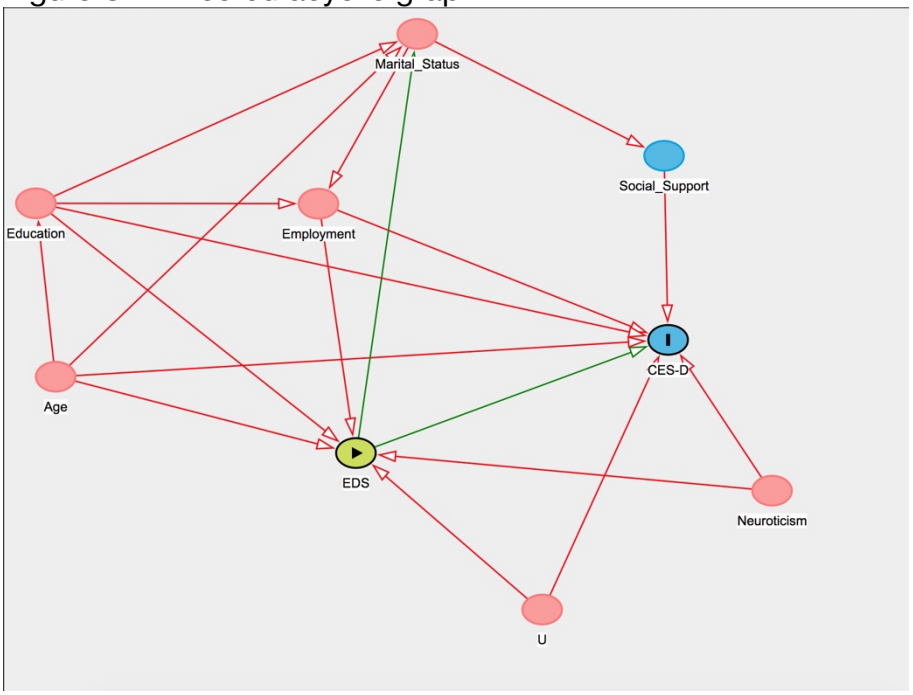
CHAPTER 1 SUPPELEMENTAL MATERIAL

Table S1. Everyday Discrimination Scale item response distribution (n(%)), African American Women's Heart & Health Study, Northern California, 2012-2013 (n=207)

	Never	Less than once a year	A few times a year	A few times a month	At least once a week	Almost everyday
You are treated with less courtesy than other people	24 (12)	25 (12)	65 (31)	40 (19)	27 (13)	26 (13)
You are treated with less respect than other people	33 (16)	30 (15)	58 (28)	32 (16)	25 (12)	29 (14)
You receive poorer service than other people at restaurants and stores	26 (13)	44 (21)	58 (28)	37 (18)	19 (9)	23 (11)
People act as if they think you are not smart	50 (24)	34 (16)	48 (23)	22 (11)	27 (13)	26 (13)
People act as if they are afraid of you	68 (33)	34 (16)	32 (16)	26 (13)	22 (11)	25 (12)
People act as if you are dishonest	65 (31)	41 (20)	37 (18)	20 (10)	15 (7)	29 (14)
People act as if they're better than you are	26 (13)	25 (12)	45 (22)	42 (20)	22 (11)	47 (23)
You are called names or insulted	96 (46)	41 (20)	33 (16)	12 (6)	12 (6)	47 (23)
You are threatened or harassed	118 (57)	32 (16)	24 (12)	10 (5)	9 (4)	14 (7)
You are followed around in stores	44 (21)	41 (20)	61 (30)	18 (9)	17 (8)	26 (13)

Question asks: "In your day-to-day life, how often. Have any of the following things happened to you because of your race, ethnicity, or skin color?"

Figure S1. Directed acyclic graph



Tables S2. Concordance and discordance in Everyday discrimination scale exposure classification by coding approach African American Women's Heart & Health Study, Northern California, 2012-2013 (n=207)

Table S2. Concordance and Discordance in EDS Exposure Classification Between Frequency- and Chronicity-Based Coding (n(%)) (n=207)

	Frequency-based coding		
Chronicity-based coding	*Low EDS*	*Mod EDS*	*High EDS*
Low EDS	59 (29)	10 (4.8)	0 (0.0)
Mod EDS	13 (6.2)	52 (25)	4 (1.9)
High EDS	0 (0)	7 (3.4)	62 (30)
n (%) discordant		34 (16)	

Table S3. Concordance and Discordance in EDS Exposure Classification Between Situation- and Frequency-Based Coding (n(%)) (n=207)

	Frequency-based coding		
Situation-based coding	*Low EDS*	*Mod EDS*	*High EDS*
Low EDS	65 (31)	17 (8.2)	3 (1.5)
Mod EDS	6 (2.9)	32 (16)	17 (8.2)
High EDS	1 (0.5)	20 (10)	46 (22)
n (%) discordant		64 (31)	

Table S4. Concordance and Discordance in EDS Exposure Classification Between Situation- and Chronicity-Based Coding (n(%)) (n=207)

	Chronicity-based coding		
Situation-based coding	*Low EDS*	*Mod EDS*	*High EDS*
Low EDS	55 (27)	25 (12)	5 (2.4)
Mod EDS	11 (5.3)	24 (12)	20 (10)
High EDS	3 (1.5)	20 (10)	44 (21)
n (%) discordant		80 (41)	

CHAPTER 3 SUPPLEMENTAL MATERIAL

Supplement A: Detailed Methodology

Search Strategy

We searched the following electronic databases with the goal of gathering studies from across academic disciplines: (1) PubMed, (2) SCOPUS, (3) PsycInfo, and (4) Sociological Abstracts.

Search terms were developed iteratively based on a preliminary review of the literature, research team expertise, content knowledge, and consultation with a public health research librarian. First, we developed preliminary search terms based on the titles and abstracts of known twelve papers examining the association between area-level racial prejudice and health outcomes [71-73,82-90]. Next, we added search terms identified by the research team and those recommended by the research librarian. We then tested preliminary search strings in multiple databases to gauge the breadth and depth of results returned. We iteratively modified search terms, string combinations, and databases to ensure all twelve known papers were identified. Once our search strategy identified all twelve known papers, we performed the formal search with no further modifications. The final set of strings were:

STRING 1: “racism” OR “stigma” OR “racial prejudice” OR “racial bias” OR “racial biases” OR “implicit racial bias” OR “explicit racial bias” OR “racial attitudes” OR “racist attitudes” OR “racial beliefs” OR “racist beliefs” OR “racial sentiment” OR “racist sentiment” OR “N-Word” OR “racial animus”

STRING 2: "project implicit" OR "general social survey" OR “Twitter” OR “Google”

STRING 3: “community-level” OR “communities” OR “county-level” OR “state-level” OR "area-level" OR “neighborhood-level” OR “regional” OR “collective”

STRING 4: “area-racism” OR “collective disrespect” OR “bias of crowds”
QUERY 1: string 1 AND string 2
QUERY 2: string 1 AND string 3
QUERY 3: string 2 AND string 3
QUERY 4: string 4

We performed our database search on April 5, 2020. One investigator entered search strings into databases 1 and 2, and another investigator entered search strings into databases 3 and 4. Our search yielded a total of 20,616 records, which were uploaded to Covidence systematic review software,[172] Two articles, published in July and September of 2020, were identified after the formal literature pull but before data extraction was complete.[74,177] We included these papers for consideration in the review

to maximize the amount of information gained from this emerging area of research. After removal of duplicates, 14,632 records proceeded to title and abstract screening. Two reviewers independently performed all screening based on inclusion and exclusion criteria. Results of the screening were compared, and disagreements were resolved via consult from a third investigator. Inclusion criteria included: (a) peer-reviewed journal article; (b) quantitative empirical study; (c) study conducted in the United States; (d) published in the English language; (e) study exposure is an indicator of bias, prejudice, animus, attitudes, sentiment, or beliefs toward or about a particular racial, ethnic, or immigrant group(s) that is measured at the individual-level and aggregated to the area-level; (f) study exposure is assessed using data from (1) surveys, questionnaires, or assessment tools, (2) social media, or (3) Google searches; and (g) study outcome is a mental or physical health outcome or health behavior.

Our title and abstract screening excluded 14,600 records, leaving 32 articles for full-text review, of which 14 met inclusion criteria. Figure 1 shows the results of these exclusions.

Data Extraction

Once the final set of included papers was identified, full-text PDFs were uploaded into MaxQDA[176] for data extraction.

We extracted standard data in accordance with PRISMA guidelines.[171] We also extracted data for our specific research questions. First, we were interested in conceptualization and framing—how were researchers thinking about area-level racial prejudice in relation to existing conceptual models for racism and health? We documented the terminology and theory used to describe the exposure, presence and content of any conceptual models, and proposed pathways to health. Second, we extracted data on empirically tested mediation and moderation of the association between area-level racial prejudice and health outcomes. In particular, we were interested in whether any association between area-level racial prejudice and health outcomes was differential by racial identity. Finally, we extracted data on key measurement and other methodological considerations.

Data Extraction Codebook

1 Background/Framing

1.1 Motivation
How are the authors motivating their approach to aggregating racial bias (e.g., to measure structural/cultural racism, to avoid self-report, some other reason, no rationale provided?)

1.2 Terminology
Terminology used to describe the exposure

1.3 Theory
Theory used? If so, which theory or theories?

1.4 Conceptual model

1.5 Pathway to health
Proposed pathway to health?

2 Study population

2.1 Exposure geography
Number of geographic units in exposure population (e.g., 208 DMAs)

2.2 Outcome pop
Number and demographic breakdown (e.g., age, sex, other) of participants in outcome assessment (e.g., outcomes on 40,000 NHW and NHB BRFSS respondents)

3 Study design

3.1 Follow-up
One time-point (cross-sectional)
Multiple time points (longitudinal
Time-to-event (survival)

3.2 Level of analysis
Ecologic - exposure and outcome measured at area-level
Multilevel - exposure at area-level (accounts for clustering), outcome at individual-level
Individual - exposure and outcome at individual-level (does not account for clustering)

3.3 Study period
Time period of data (exposure, outcome, and covariates)

4 Study setting

4.1 Study area
E.g., California, US, global

4.2 Exposure scale
At what geographic scale was the exposure measured?

4.3 Outcome scale
At what geographic scale was the outcome measured?

4.4 Covariate scale
At what geographic scale were covariates measured

5 Exposure

5.1 Exposure(s)

5.2 Data source
Project Implicit
Google
General Social Survey
Twitter
Other

5.3 Number aggregated
Number of individual observations aggregated (e.g., n=1 million IAT responses were aggregated to the county-level) (if reported)

5.4 Specification
Implicit or explicit racial bias data
Restrictions
Weighting
Google search terms queried
Continuous or binary
Coding, cutpoints, etc.
Any information on validity (either based on prior literature, or tested in the study)

6 Outcomes

6.1 Outcome(s)
What was the primary study outcome?

6.2 Data source
What was the data source for the study outcome?

6.3 Assessment
E.g., self-report, biomarker, administrative records

6.4 Specification
E.g., Continuous, binary coding/cutpoints used; other details

7 Confounder adjustment

7.1 How identified?

How were confounders identified (e.g., literature review, DAG, data-driven approaches)?

7.2 Area-level
What area-level confounders were identified and how were they measured?

7.3 Individual-level
What individual-level confounders were identified and how were they measured?

7.4 How addressed?
How was confounding addressed (e.g., multivariable regression, propensity score matching, econometric models?)?

8 Findings

8.1 Statistical model
Statistical model used and any modeling notes (e.g., robust SEs, sensitivity analyses performed, etc)

8.2 MOA
Measure of association and 95% confidence interval for main results, subgroup effects, and any sensitivity analyses

8.3 Findings
Tag to highlight summary of findings

9 Mechanisms

9.1 Area-level
Area-level mediation or effect measure modification (formal interaction or stratified results)

9.2 Individual-level
Individual-level mediation or effect measure modification (formal interaction or stratified results)

9.3 Differential?
Association differential or non-differential by racial identity (assessed via formal interaction or race-stratified results)?

10 Limitations

10.1 Investigator
Limitations identified by the investigator

10.2 Research team

Limitations identified by the research team

11 Notes

11.1 Implications

Implications for future research

11.2 Other refs

Any other references to include in the review that we missed in our literature pull

11.3 Notable

Anything else you find notable or want to come back to; memorable quotes

Table S1A. Overview of studies

Source	Study design		Study sample	Study setting				
First author, year	Level of analysis	Follow-up	Study sample (descriptives if provided)	Study area	Years of data for Exposure, Outcome, Covariates	Exposure scale	Outcome scale	Covariate scale
Kennedy et al., 1995	Ecologic	Cross-sectional	N/A (rates)	39 US states (not specified which states)	E: 1986-1990 O: 1990 C: 1990	Individual, aggregated to state	State	State
Lee et al., 2015	Multilevel	Prospective (discrete-time event history)	n=10,950 Black and white GSS respondents across 100 PSUs; Mage=45 years 55% female, 85.7% white, 14.3% = Black	US	E: 1993-2002 O: 1993-2008 C: 1990-2002	Individual, aggregated to PSU	Individual	Individual and PSU
Morey et al., 2018	Multilevel	Prospective (survival)	n=13,242 immigrant GSS respondents across 123 PSUs (Mage=43.5, 53% female, 79% white, 14% Black, 8% Other Race)	US	E: 1993-2010 O: 1993-2014 C: 1993-2014	Individual, aggregated to PSU	Individual	Individual and PSU
Leitner et al., 2016a	Ecologic	Cross-sectional	Study 1: n=199,159 Black and white BRFSS respondents (11.8% Black, 88.9% white) but outcomes were modeled as rates Study 2: NA (rates)	US	E: 2003-2013 O1: 2012 O2: 2003-2013 C: 2000, 2005-2013	Test, aggregated to county	Study 1: County Study 2: County	Study 1: County Study 2: County
Leitner et al., 2016b	Ecologic	Cross-sectional	N/A (rates) Black death rate per 100,000: M = 352.595, SD = 84.806; White death rate per 100,000: M = 270.477, SD = 54.2	US	E: 2003-2013 O: 2003-2013 C: 2005-2013	Test, aggregated to county	County	County

Orchard & Price, 2017	Multilevel	Cross-sectional	n=31.5 million births (white Mage = 27.78, SD = 6.04, 15% finished college; Black Mage = 25.84, SD = 6.22, 8% finished college)	US	E: 2002-2012 O: 2002-2012 C: 2002-2013	Test, aggregated to county	Individual	Individual and county
Hehman et al., 2017	Ecologic	Cross-sectional	n=875 individuals confirmed as killed by police officers in the United States (Mage = 37.3 years, SD = 13.3; 4% female)	US	E: 2003-2013 O: 1/1/15-9/30/15 C: 2010-2013	Test, aggregated to CBSA	CBSA	CBSA
Chae et al., 2015	Ecologic	Cross-sectional	23.1 million person-years across 196 DMAs (49.3% aged 45+; 52.81% female) but outcomes were modeled as rates	US (except AK)	E: 2004–2007 O: 2004–2009 C: 2000, 2004-2009	Search, aggregated to DMA	DMA	DMA
Chae et al., 2018	Multilevel	Cross-sectional	n=2,332,216 births to Black women across 196 DMAs (maternal age: 6.3% <18, 83.6% 18-34, 10.1% 35+)	US (except AK)	E: 2004-2007 O: 2005-2008 C: 2005-2010	Search, aggregated DMA	Individual	Individual and DMA
McKetta et al., 2018	Multilevel	Prospective (survival)	N=16,580 Black and white PSID respondents (66.1% white, 33.9% Black)	US (except AK)	E: 2004-2007 O: 1990-2009 C: 1990	DMA, aggregated to state	Individual	Individual and state
Nguyen et al., 2018	Multilevel	Cross-sectional	n=3,988,733 births (birthing persons – 53% were white, non-Hispanic, and 77% were U.S. born)	Contiguous US + DC	E: March 2015–April 2016 O: 2015 C: 2015	Tweet, aggregated to state	Individual	Individual and state
Huang et al., 2020	Individual	Cross-sectional	n=450,016 participants (range: n=433,434 to n=433,680 across outcomes)	Contiguous US + DC	E: 2015-2018 O: 2017 C: 2017	Tweet, aggregated to state	Individual	Individual

Nguyen et al., 2020	Multilevel	Serial cross-sectional	N=9,988,030 for gestational age, n=9,985,402 for birth weight (birthing persons – Mage= 29 years, 59.74% married, 85.99% completed at least high school)	Contiguous US + DC	E: June 2015-December 2017 O: 2015-2017 C: 2013-2017	Tweet, aggregated to state	Individual	Individual and state
Hswen et al., 2020	Ecologic	Time-series	n=8,314 Hispanic Gallup respondents (no descriptives provided)	US	E: 8/29/2016-1/16/2017 O: 8/29/2016-1/16/2017	Tweet, aggregated to US	Individual, aggregated to US	NA

Table S1B. Study measures

Source	Exposure				Outcome			
First author, year	Exposure	Data source	# aggregated	Operationalization/specification	Outcome	Data source	Assessment	Operationalization/specification
Kennedy et al., 1995	Collective disrespect	GSS	n=7,679	Individual-level questions: "On average, blacks have worse jobs, income, and housing. Do you think the differences are due to (a) discrimination, (b) less in-born ability to learn, (c) lack of chance for education that it takes to rise out of poverty, (d) less motivation or willpower to pull themselves out of poverty?* * Each item was dichotomized separately Aggregate measure: state-level % of respondents who answered in the affirmative to each item Weighting: Post-stratification weights based on age, race, educational attainment Specification: continuous	Age-adjusted all-cause Black and white mortality rates	NCHS death records	Administrative (death) records	Measure: Directly age-standardized to the US population of Blacks and whites, and expressed as the number of deaths per 100,000 persons. Specification: rate per 100,000
Lee et al., 2015	Community-level racial prejudice	GSS-NDI	n=13,355 (14,513 GSS respondents-1,158 with missing racial prejudice data)	Individual-level questions: 1. "On the average, negroes/blacks/African-Americans have worse jobs, income, and housing than white people. Do you think these differences are caused by the fact that most negroes/blacks/African-Americans have less in-born ability to learn?" 2. "Do you think these differences are because most	All-cause mortality (survival)	GSS-NDI	Administrative (death) records	Measure: 0=alive in 2008, 1=died by 2008 Specification: binary (survival)

				negroes/blacks/African-Americans just don't have the motivation or willpower to pull themselves up out of poverty?" 3. "Do blacks tend to be unintelligent or tend to be intelligent?" (and "Do whites tend to be unintelligent or tend to be intelligent?") 4. "Do blacks tend to be hard working or lazy?" (and "Do whites tend to be hard working or lazy?") 5. "Do you think there should be laws against marriages between Negroes/Blacks/African-Americans and whites?"* * Each item was dichotomized, then averaged across items <u>Aggregate measure:</u> PSU-level average scores <u>Specification:</u> Standardized & centered (continuous)				
Morey et al., 2018	Community-level anti-immigrant prejudice	GSS-NDI	n=2,427	<u>Individual-level questions:</u> 1. "Do you think the number of immigrants to America nowadays should be increased a lot, increased a little, remain the same as it is, reduced a little, or reduced a lot?" 2. Respondents were asked how much they agreed or disagreed with the following four statements: (1) "America should take stronger measures to exclude illegal immigrants," (2) "Immigrants take jobs away from people who were	All-cause mortality (survival)	GSS-NDI	Administrative (death) records	<u>Measure:</u> 0=alive in 2014 1=died by 2014 + censored amount of time at risk over the study period <u>Specification</u>: binary (survival) and continuous (time-to-event)

				born in America," (3) "Immigrants increase crime rates," and (4) "Immigrants are generally good for America's economy." Responses were coded on a five-point Likert scale ranging from "agree strongly" to "disagree strongly."* * Each item was dichotomized, then summed across items Aggregate measure: PSU-level average scores Specification: Continuous and dichotomous (+/- 1SD from the mean)				
Leitner et al., 2016a	white county-level racial bias	Project Implicit Race IAT	n=1,391,632 white IAT responses	Individual implicit measure: keyboard association test with D-score Individual explicit measure: temperature difference Aggregate measure: county-level average implicit and explicit scores of white IAT respondents Weighting: Post-stratification weights based on age Specification: Continuous and dichotomous (+/- 1SD from the mean)	Study 1: Black and white circulatory-disease risk (% without access to health care*, % with circulatory disease) Study 2: Black and white age-adjusted circulatory disease mortality rates * % without access to health care does not meet inclusion criteria for the review.	Study 1: BRFSS Study 2: NCHS death records	Study 1: self-report (telephone interview) Study 2: administrative (death) records	Study 1: circulatory disease risk Circulatory disease diagnosis question "Has a doctor, nurse, or other health professional ever told you that you had a heart attack, also called a myocardial infarction?" or "…angina or coronary heart disease?" Coding: 0="no," 1="yes" Aggregation: averaged at the county level to calculate county % without healthcare access and % with either diagnosis. Specification: prevalence (continuous), examined separately and as a B-W difference Study 2: circulatory disease mortality

								Measure: Black and white deaths from circulatory diseases (e.g., heart disease; Internal Statistical Classification of Diseases and Related Health Problems codes I00–I99). Age adjusted based on 2000 standard population (for each racial group) Specification: rates per 100,000, examined separately and as B-W difference
Leitner et al., 2016b	Black and white county-level ingroup bias	Project Implicit Race IAT	n=250,665 Black IAT responses, n=1,391,632 white IAT responses	Ingroup bias: white respondents' pro-white/anti-Black bias and Black respondents' pro-Black/anti-white bias (i.e., ingroup favoritism) Individual implicit measure: keyboard association test with D-score, scaled for ingroup Individual explicit measure: temperature difference, scaled for ingroup Aggregate measure: county-level average implicit and explicit scores Weighting: Post-stratification weights based on age Specification: Continuous and dichotomous (+/- 1SD from the mean)	Age-adjusted Black and white circulatory disease mortality rates	NCHS death records	Administrative (death) records	Measure: Black and white deaths from circulatory diseases (e.g., heart disease; Internal Statistical Classification of Diseases and Related Health Problems codes I00–I99). Age adjusted based on 2000 standard population (for each racial group) Specification: rate per 100,000
Orchard & Price, 2017	Community-level racial prejudice	Project Implicit Race IAT	n=1.8 million IAT responses aged 18+ (mean age=28,	Individual implicit measure: keyboard association test with D-score Individual explicit measure: preference measure	Black and white rates of adverse birth outcomes	NCHS birth records	Administrative (birth) records	Measure: Binary PTB: gestational age < 37 weeks Binary LBW: < 2500 g Specification: Black and white rates per 1,000 births

			and 59% women)	Aggregate measure: county-level average implicit and explicit scores Weighting: Post-stratification weights based on age + gender* Specification: Standardized (continuous) and dichotomous (+/- 1SD from the mean)				(B-W difference assessed with interaction term)
Hehman et al., 2017	Regional racial biases of residents (Black-white bias and weapons stereotype)	Project Implicit Race IAT and Weapons IAT	n=1,860,818 Black and white Race IAT responses, n=295,235 Black and white Weapons IAT responses	Individual implicit measure: keyboard association test with D-score (race IAT and weapons IAT) Individual explicit measure: temperature difference (race IAT) Aggregate measure: CBSA-level average implicit and explicit scores, separately for Black and white respondents Weighting: No Specification: Untransformed (continuous)	Disproportionate lethal force against Black and white people relative to their population shares	The Guardian police killing database	Traditional reporting with police reports Fact-checked witness statements; monitoring of regional news; other open-sourced police fatality databases	Measure: % of Black people living in each CBSA was subtracted from the % of Black people killed in each CBSA relative to the total amount of individuals killed by police officers. Higher score on this variable reflected greater usage of lethal force with Black people than would be expected based on the CBSA population (i.e., disproportionate lethal force). An identical score was calculated for NH white people Specification: continuous
Chae et al., 2015	Area racism	Google Trends data compiled by SSD (2014)	NA	Aggregate measure: DMA-level proportion of total Google searches containing the "n-word." (singular or plural, ending in "-er(s)" but not "-a(s)") Specification: Standardized (continuous)	Age-adjusted Black all-cause and cause-specific (heart disease, cancer, stroke, and diabetes) mortality rates	NCHS death records	Administrative (death) records	Measure: Black mortality rates weighted using the US 2000 standard population were calculated for all-cause mortality and the four leading specific causes of death among Black people identified using International Classification of Disease, Version 10 codes: heart disease (I00-I09, I11, I13, I20-I51); cancer (C00-C97); stroke (I60-I69); and diabetes (E11-E14)

								Specification: rate per 100,000 person-years
Chae et al., 2018	Area racism	Google Trends data compiled by SSD (2014)	NA	Aggregate measure: DMA-level proportion of total Google searches containing the "n-word." (singular or plural, ending in "-er(s)" but not "-a(s)") Specification: Standardized (continuous)	PTB and LBW among NH Black women	NCHS birth records	Administrative (birth) records	Measure: PTB: gestational age < 37 weeks LBW: < 2500 g Specification: Binary
McKetta et al., 2018	State-level racial animus	Google Trends data compiled by SSD (2014)	NA	Aggregate measure: State-level proportion of total Google searches containing the "n-word." (singular or plural, ending in "-er(s)" but not "-a(s)") (DMAs aggregated to state-level) Specification: Quartiles	Black and white SRH and Black-white differences in SRH (also movement across states)	PSID	Self-report (telephone interview)	Measure: At each interview wave, respondents were asked to report whether their health was "excellent, very good, good, fair, or poor." Poor SRH if respondent self-rated poor or fair health (vs excellent or very good) in at least two consecutive interviews Specification: Binary (survival)
Nguyen et al., 2018	Twitter-derived sentiment toward racial and ethnic minorities	Twitter API	n=1,249,653 tweets containing at least one race-related term	Sample: random 1% of geotagged Tweets from Twitter's API (March 2015–April 2016), subset Tweets referencing racial or ethnic groups/slurs using one or more of 398 race-related keywords Sentiment analysis: identified Tweets referencing black, Hispanic, Asian, white, and Middle Eastern groups and used machine learning algorithm with hand-coded training data to classify sentiment of Tweets: 1=positive, 0=negative/neutral Aggregate measure: state-level % of Tweets that was positive (racial minorities	LBW, VLBW, and PTB among birthing persons of various racial/ethnic groups	NCHS birth records	Administrative (birth) records	Measure: LBW: ≤ 2499g VLBW: ≤ 1499g PTB: gestational age < 37 weeks based on the obstetric estimate of gestation at delivery (OE). Specification: Binary for each outcome

				overall, and broken out by racial group) Specification: tertiles (ref=T3)				
Huang et al., 2020	Twitter-characterized sentiment toward racial and ethnic minorities	Twitter API	n=30,977,747 tweets containing at least one race-related term	Sample: random 1% of geotagged/place-labeled Tweets from Twitter's API (2015-2018), subset Tweets referencing racial or ethnic groups/slurs using one or more of 518 race-related terms Sentiment analysis: used machine learning algorithm with hand-coded training data to classify sentiment of Tweets: negative (1=negative, 0=positive/neutral) and positive (1=positive, 0=negative/neutral) Aggregate measure: state-level % of Tweets that was negative and % that was positive Specification: Tertiles (ref=T1 for both)	CVD outcomes (e.g., hypertension, stroke) among various racial/ethnic groups	BRFSS	Self-report (telephone interview)	Question: Has a doctor, nurse or other health professional ever told you that you had … …hypertension, diabetes, obesity, stroke, myocardial infarction (MI), coronary heart disease (CHD)? Measure: Each outcome coded as binary (0=no, 1=yes). Any CVD if they answered "yes" to one or more. BMI: ≥30 kg/m2 was defined as obesity. Specification: Binary for each outcome
Nguyen et al., 2020	State-Level Racial Attitudes Assessed From Twitter Data	Twitter API	n=26,027,740 tweets from 2,498,717 Twitter users containing at least one race-related term	Sample: random 1% of geotagged/place-labeled Tweets from Twitter's API (June 2015-Dec 2017), subset Tweets referencing racial or ethnic groups/slurs using one or more of 518 race-related terms Sentiment analysis: identified Tweets referencing black, Hispanic, Asian, white, and Middle Eastern groups and used machine learning algorithm with hand-coded training data to classify sentiment of	LBW and PTB among birthing persons of various racial/ethnic groups	NCHS birth records	Administrative data	Measure: LBW: ≤ 2499 g. PTB: gestational age < 37 weeks based on the obstetric estimate of gestation at delivery (OE) Specification: Binary

				Tweets: negative (1=negative, 0=positive/neutral) and positive (1=positive, 0=negative/neutral) Aggregate measure: state-level % of Tweets that was negative and % that was positive Specification: Tertiles (ref=T1 for both)				
Hswen et al., 2020	Negative sentiment towards Mexicans and Hispanics during the 2016 presidential election	Twitter licensing agreement	n=2,809,641 tweets from 943,766 users containing terms Mexican(s) and/or Hispanic(s) (1,594,845 retweets)	Sample: full stream of tweets from Twitter over a 20-week period: 10 weeks before and 10 weeks after the 2016 United States presidential election Sentiment analysis: identified Tweets referencing Mexican(s) or Hispanic(s) (with and without #) and used VADER method to assign Tweets a continuous sentiment score ranging from -1 (most negative) to 1 (most positive), also collapsed into negative (< -0.5), positive (> +0.5), or neutral (-0.5 to +0.5). Aggregate measure: population-level weekly averages (whole US) Specification: weekly mean score and % negative, positive, and neutral	Daily negative mental wellbeing (worry)	Gallup-Sharecare Well-Being Index	Self-report (telephone interview)	Measure: Emotional well-being index measures Americans' daily experiences, and respondents categorize their responses as thriving, struggling, or suffering in the areas that measure wellbeing Specification: population-level weekly average % worry

Table S1C. Estimation and results

Source	Confounders				Estimation		Mediators/moderators evaluated		
First author, year	How identified	Area-level	Individual-level	How controlled	Statistical model	Adjusted MOA (95% CI or SE)	Mediators/moderators	How assessed	Mediation/moderation findings
Kennedy et al., 1995	Cited literature & defined as potential confounders: "some evidence suggests that low income and poverty are linked to depletion in social capital. Since income levels and poverty are also potential predictors of mortality, we evaluated these variables as potential confounders in the relationship between collective disrespect and mortality."	Median income, % in poverty	Accounted for age in the creation of rates	MV regression	OLS regression	<u>Black mortality:</u> No ability: Beta=336.5, SE=93.4, p=0.0009 No willpower: Beta=256.1, SE=83.6, p=0.004 Discrimination: Beta=-290.1, SE=99.0, p=0.006 Lack of educational opportunity: Beta=-246.9, SE=83.9, p=0.006 <u>White mortality:</u> No ability: Beta=182.4, SE=71.9, p=0.01 No willpower: Beta=148.5, SE=62.5, p=0.02 Discrimination: Beta=-147.1,	Race (Black or white)	Examined race-specific mortality rates	Collective disrespect was associated with Black and white mortality rates but <u>results were stronger for Black mortality (>10% difference)</u>

						SE=75.0, p=0.06 Lack of educational opportunity: Beta=-173.8, SE=60.3, p=0.007 *Betas for one-unit change in collective disrespect*			
Lee et al., 2015	Data-driven: All PSU-level covariates were chosen because they were significantly correlated with racial prejudice in bivariate models and there- fore could be potential confounders of the relationship between racial prejudice and mortality Formation of candidate confounder list not specified	average number of people living below the federal poverty line (adjusted for family size and survey year), median income, average years of educational attainment, % Black, located in the South, dissimilarity index, political affiliation index	Race (white, Black), gender[+], age at the time of the interview, marital status, household income, educational attainment	MV regressi on	3-level HLM survival model	<u>Community-level racial prejudice, adjusting for individual-level prejudice and confounders:</u> OR=1.24; 95% CI=1.04, 1.49 *OR for 1SD change in community-level racial prejudice*	Moderators: Individual race (Black or white), Individual-level prejudice Mediator: community-level social capital	Multiplicative interaction terms in regression model: 1. race*individual-level prejudice; race*community-level prejudice 2. individual-level prejudice *community-level prejudice Mediation: change-in-estimate approach	<u>Race did not moderate</u> the association between community-level prejudice on mortality (race*prejudice interactions ns) or mediation through community-level social capital (described below) There was a significant interaction between <u>Individual*community-level prejudice</u>: OR = 0.74; 95% CI = 0.58, 0.95, indicating that individuals low in racial prejudice but living in higher-prejudice communities had

									the highest level of mortality risk. Mediation: Social capital was inversely related to community-level prejudice (r =–0.41; P <.01), indicating that communities with higher levels of prejudice had lower levels of social capital. When social capital was controlled in the fully adjusted model, PSU-level racial prejudice was no longer significantly associated with mortality.
Morey et al., 2018	Prior research: We included variables that prior research suggested may be potential confounders of the association between anti-immigrant prejudice and mortality.	% foreign-born, mean years of education, mean family income, % who identify as politically conservative, survey year for anti-immigrant score	Gender,+ age, marital status , years of education, unemployment, family income, self-rated health at baseline	MV regression	Cox proportional hazards models with clustered SEs	Community-level anti-immigrant prejudice and mortality main effects *ns*: HR=1.05 95% CI=0.93, 1.19 *HR for 1-unit change in anti-immigrant prejudice score*	Nativity status, Race (Black, white, Other – *sensitivity analysis restricted "other race" to Asian and Hispanic)*	Multiplicative interaction terms in regression model: race*community-prejudice; nativity*community-prejudice; race*nativity*community-prejudice; Also stratified results by race (Table 3)	Race*nativity moderated nativity*prejudice: ns race*prejudice: ns race*nativity*prejudice: sig (F-test=4.04, p=0.018) – interpretation: *the association between anti-immigrant prejudice and mortality for US-born respondents was significantly different compared to*

									foreign-born respondents <u>Stratified findings by race and nativity:</u> <u>US-born "other race":</u> The mortality hazard ratio for US-born respondents living in high-prejudice communities (HR=2.63 [95% CI: 0.53, 13.12]) was 171% higher than US-born respondents living in low-prejudice communities (HR=1.54 [95% CI: 0.75, 3.18]). <u>Foreign-born "other race":</u> The mortality hazard ratio for foreign-born respondents living in the high-prejudice communities (HR=0.15 [95% CI: 0.02, 1.20]) was 287% lower than foreign-born respondents living in low-prejudice communities (HR=0.43 [95% CI: 0.17, 1.09]). *HR comparing mortality in high*

									(1SD above the mean) vs low (1SD below the mean) prejudice communities Results restricted to Asian and Hispanic "other race" respondents showed similar patterns but were less precise due to small number of respondents.
Leitner et al., 2016a	<u>Not stated with two exceptions:</u> (1) Geomobility: "Importantly, a relationship between Blacks' racial bias and ingroup health could be driven by social selection forces." (2) Age bias: "To examine whether any effects were specific to racial bias, or generalized to bias on nonracial dimensions"	<u>Study 1:</u> total population, Black-to-white ratio, dissimilarity index of segregation, Black geographic mobility, housing density, urbanicity (number of housing units per square mile), implicit and explicit age bias, and average of Black and white: high school graduation rates, MHI	<u>Study 1:</u> accounted for sex, age, and race in creation of rates <u>Study 2:</u> accounted for age and race in creation of rates	MV regressi on	GEE with robust standard errors and simple slopes	<u>White explicit bias and circulatory disease death rates:</u> <u>Black death rates</u> (positive, stronger): b=43.20, SE=12.10, p=0.0004, <u>White death rates</u> (positive, weaker): b=13.90, SE=4.97, p=0.005 Implicit bias ns (simple slopes estimates not shown) *b for 1-point*	Effect modification: Race (Black or white) Mediation: Black-white disparities in health behaviors (smoking, drinking, and exercise)	Multiplicative interaction term in regression model: Race*white implicit bias Race*white explicit bias (also 3-way interaction with race*sex, but results were ns) Sig interaction effects explored via simple slopes analysis Mediation: Change-in-estimates approach	<u>Study 1</u> – NA (no main effects on circulatory disease diagnosis) <u>Study 2</u> – <u>Race moderated</u> association between explicit racial bias and healthcare access (sig race*implicit bias interaction). <u>Simple slopes:</u> whites' explicit racial bias was associated with white and Black circulatory disease death rates, but stronger association with Black rate (race*implicit ns) <u>Mediation</u>: Black-white disparities

		past 12 months, unemployment, % in poverty *each covariate interacted with individual race <u>Study 2</u>: same as study 1 + neoplasm (cancer) death rate				*increase in racial bias of white people*			in <u>health behaviors did not mediate</u> the relationship between explicit/implicit bias and death rate disparity.
Leitner et al., 2016b	"We adopted an analytic approach that could test whether Blacks' bias remained a predictor of Blacks' death rate when we controlled for a large set of socio-demographic characteristics and whites' biases in the same county." Further explicated	Total population, Black population, Black/white MHI, Black/white high school graduation rate, Black/white poverty rate, Black/white unemployment rate, dissimilarity index of segregation, housing density, Black geographic mobility, income inequality,	NA	MV regression	Analysis 1: GEE with robust SEs and simple slopes	<u>Black ingroup bias and Black death rate:</u> *Explicit*: b=0.005, SE=6.20, Beta <0.001, p=0.99 *Implicit:* b=157.24, SE=34.04, Beta =0.49, p<0.0001 <u>White ingroup bias and white death rate:</u> *Explicit:* b=19.04, SE=4.98, p=0.0001	Race (Black or white)	Multiplicative interaction term in regression model: Race*ingroup implicit bias Race*ingroup explicit bias (also looked at higher order interactions with ingroup implicit*explicit*race, but results were ns) Sig interaction effects explored via simple slopes analysis	<u>Race moderated</u> the association between implicit and explicit racial bias and ingroup death rates: Implicit ingroup bias was associated with Black but not white death rates Explicit ingroup bias was associated with white but not Black death rates

	rationale for sex ratio (previous research), income inequality (previous research), and geomobility (conjecture, content knowledge)	Black/white male-to-female ratio, ingroup racial bias of other group (e.g., models for Black bias and health controlled for whites' implicit and explicit ingroup bias)				*Implicit*: b=23.81, SE=28.10, p=0.40 *b for 1-point increase in ingroup bias Estimates derived from simple slope analysis with race*ingroup bias interactions*			
Orchard & Price, 2017	"We include additional covariates to reduce the possibility of county-level prejudice being correlated with other individual and county characteristics."	Total population, unemployment rate, % college graduates, % Black*, Black poverty rate*, sexual orientation IAT, gender-career IAT * interacted with birthing person's race	Maternal age, marital status, education, and 17 different pregnancy risk factors (e.g., high blood pressure, previous preterm birth, etc.); child gender+ and birth order of the child.	MV regression	Weighted least squares regression with clustered SEs and state- and year- fixed effects	<u>*Implicit*</u> LBW: The black-white gap in low birth weight is 14% larger in counties with high vs low implicit racial prejudice. PTB: The black-white gap in low birth weight is 29% larger in counties with high vs low implicit racial prejudice. <u>*Explicit:*</u>	Birthing person's race (Black or white); County of residence vs. county of birth	Multiplicative interaction term in regression model: county prejudice*birthing person's race. Used models to estimate race-specific effects, and plotted stratified results. Also stratified results on bias in county of residence vs. birth county	<u>Race moderated</u> the association between community-level prejudice and birth outcomes (sig interaction). Findings showed stronger associations among Black birthing persons and no (or even protective) associations among white birthing persons. Prejudice in community of birth was more strongly associated with birth outcomes than prejudice in

						LBW: The black-white gap in low birth weight is 22% larger in counties with high vs low explicit racial prejudice. PTB: The black-white gap in low birth weight is 36% larger in counties with high vs low explicit racial prejudice. <u>Other:</u> (1) When implicit and explicit bias were modeled together, only explicit remained significant predictor of B-W birth outcome gaps. (2) Explicit prejudice in county of birth more strongly associated with B-W			community of residence.

						birth outcome gaps than county of residence (results similar for implicit). (3) Results unique to racial bias (no findings for gender-career or sexual orientation bias)			
Hehman et al., 2017	Inflated model used to "develop initial predictive models of lethal force" (supplemental material include more parsimonious models determined using data-driven approach: forward and backward stepwise regression)	Black/white MHI, Black/white % with HS or equivalent degree, isolation index of segregation, violent crime, unemployment, population density, and total (race disaggregated) lethal force	NA	MV regression	Linear regression	<u>Model with race-IAT</u> White *implicit*: b=4.13, SE=1.90, p=0.031 White *explicit*: b=-0.52, SE=0.29, p=0.079 Black *implicit*: b=-1.13, SE=0.84, p=0.182 Black *explicit*: b=0.12, SE=0.14, p=0.40 <u>Model with race-IAT and</u>	NA	NA	NA, but note they did calculate a disproportionate lethal force measure for white people and found they were not being killed disproportionately. Therefore, estimation was just for the association between regional racial bias and disproportionate killing of Black people.

						<u>weapons-IAT</u> Implicit/explicit racial bias of whites and Blacks – not significant White *implicit threat stereotypes*: b=5.50, SE=1.63, p=0.001 *b for 1-point increase in race- and weapons-IAT of white people*			
Chae et al., 2015	Adjusted for "relevant area-level covariates"	% in urbanized area (>50,000 population), % Black, % Blacks with up to a high school education, % Black households in poverty, white mortality rate	Accounted for age group, sex, year of death, census region in creation of rates	MV regression	Negative binomial regression model with Huber-white clustered SEs	<u>All cause</u>: MRR=1.04, 95% CI = 1.02, 1.06 <u>Heart disease</u>: MRR=1.04, 95% CI=1.02, 1.07 <u>Cancer</u>: MRR=1.03, 95% CI=1.00, 1.05 <u>Stroke</u>: MRR=1.03, 95% CI=1.00,	NA	NA	NA

						1.07 <u>Diabetes</u>: MRR=0.95, 95% CI=0.88, 1.019 *MRR for 1SD increase Google searches for N-word*			
Chae et al., 2018	"conceptual relevance" + data-driven (changes-in-estimates) discussed	Census region, % Black, % in urbanized area (>50,000 population), % Black w/ <HS degree or equivalent, % Black in poverty	Maternal age	MV regression	Log-binomial regression model fit with GEE	PTB: PR=1.05, 95% CI=1.02, 1.09, LBW: PR=1.05, 95% CI=1.02, 1.07 *PR for 1SD increase in Google searches for N-word*	NA	NA	NA
McKetta et al., 2018	Adjusted for "relevant confounders"	Median income, % Black population (sensitivity), Google searches for N-word in SRH->movement model	At baseline: SRH, age, education level	MV regression	Incident SRH: Cox PH Movement: Logistic regression	<u>Incident SRH among White respondents:</u> Q2: HR=1.19, 95% CI=1.07, 1.32 Q3: HR=1.13, 95% CI=1.04, 1.22	Race (Black or white)	Multiplicative interaction term in regression model: race*state-level racial animus	Race <u>did not moderate</u> the association between state-level racial animus and poor SRH (interaction term ns)

						Q4: HR=1.33, 95% CI=1.20, 1.47 <u>Incident SRH among Black respondents:</u> Q2: HR=1.43, 95% CI=1.12, 1.82 Q3: HR=1.31, 95% CI=1.05, 1.63 Q4: HR=1.20, 95% CI=0.95, 1.50 *Ref = Q1 racial animus (Google searches for N-word)*			
Nguyen et al., 2018	<u>Individual</u>: "to adjust for potential confounding of the relationship between neighborhood environments and birth	MHI, % NH white	Maternal age, marital status, race, Hispanic ethnicity, education, BMI, smoking status during pregnancy,	MV regression	Log Poisson regression models with robust SEs	<u>T1 vs T3 positive sentiment toward racial/ethnic minorities</u> <u>LBW:</u> PR=1.06, 95% CI=1.04, 1.07	Birthing person's race/ethnicity (white vs Hispanic or nonwhite or foreign-born)	Stratified subgroup analyses (did not test for statistical interaction)	Race/ethnicity <u>did not moderate</u> association between Twitter sentiment and birth outcomes: Results from subgroup analyses restricted to racial/ethnic

	outcomes." <u>State</u>: "to account for between-state differences in compositional characteristics."		first birth indicator, prenatal care in the 1st trimester indicator			<u>VLBW:</u> PR=1.09, 95% CI=1.06, 1.12 <u>PTB:</u> PR=1.10, 95% CI=1.10, 1.11 <u>Note</u>: sentiment towards specific racial/ethnic groups showed a similar pattern of results			minorized birthing persons did not differ substantially from those seen for the full population of birthing persons (differences in PRs <10%).
Huang et al., 2020	Not stated	% NH white, % NH Black, % Hispanic, MHI	Age, sex, education, race/ethnicity, and marital status	MV regression	Poisson regression	*<u>T3 vs T1 negative sentiment toward racial/ethnic minorities</u>* <u>Hypertension</u>: PR=1.11, 95% CI=1.08, 1.14 <u>Diabetes</u>: PR=1.15, 95% CI=1.08, 1.22 <u>Obesity</u>: PR=1.14, 95% CI=1.10, 1.18	gender+ and race/ethnicity	Assessed statistical interactions: Sentiment*sex Sentiment*race/ethnicity	<u>Race and sex did moderate</u>, but findings depended on the outcome: + In general, effects were stronger for women (except on diabetes and obesity) + Negative sentiment and hypertension, MI, and any CVD = stronger for non-Hispanic whites and non-Hispanic blacks than other race/ethnicity groups

						Stroke: PR=1.30, 95% CI=1.16, 1.46 MI: PR=1.14, 95% CI=1.03, 1.25 CHD: PR=1.09, 95% CI=1.00, 1.19 Any CVD: PR=1.16, 95% C=1.09, 1.24 *T3 vs T1 positive sentiment toward racial/ethnic minorities* Hypertension: PR=0.97, 95% CI 0.94, 1.00 Diabetes: PR=0.94, 95% CI 0.90, 0.99 Obesity: PR=0.97, 95% CI 0.94, 1.00 Stroke: PR=0.89, 95% CI 0.80, 0.98			+ Negative sentiment and diabetes, obesity, and stroke = stronger in Hispanics than any other racial/ethnic groups + Positive sentiment and hypertension, diabetes, and obesity = effects more protective in non-Hispanic blacks than non-Hispanic whites

						MI: PR=0.91, 95% CI 0.83, 0.98 CHD: PR=0.94, 95% CI 0.86, 1.02 Any CVD: PR=0.90, 95% CI 0.86, 0.95			
Nguyen et al., 2020	Individual: "We adjusted for potential confounders of the association between racial sentiment and birth outcomes." State: "to account for state-level compositional differences in demographic and economic characteristics."	% NH Black, % Hispanic, population density, Southern state indicator, economic disadvantage composite. (% unemployed; % some college education, % high school diploma, % children in poverty, % single parent household, MHI)	Maternal age, marital status, race, Hispanic ethnicity, education, BMI, smoking status during pregnancy, first birth indicator, prenatal care in the 1st trimester indicator, birth year	MV regression	log binomial regression models with clustered SEs	LBW: T2: IR=1.08, 95% CI=1.03-1.13 T3: IR=1.08, 95% CI=1.04-1.13 PTB: T2: IR=1.09, 95% CI=1.04-1.13 T3: IR=1.08, 95% CI=1.00-1.14 *Ref=T1 negative sentiment toward racial/ethnic minoritized persons*	Birthing person's race (Black NH, white NH, Asian NH, Hispanic, and all minoritized persons)	Stratified subgroup analyses (did not test for statistical interaction)	Race did not moderate: State-level sentiment toward all minoritized people was associated with adverse birth outcomes among all birthing persons (differences in IRs <10%). Negative sentiment toward racial/ethnic minoritized persons (T3 vs T1 (ref)): *Among all racial/ethnic minoritized birthing persons:* LBW: IR=1.13 (1.06-1.21) PTB: IR=1.10 (1.05-1.16) *Among White birthing persons:* LBW: IR=1.08

									(1.03-1.14) PTB: IR=1.08 (1.00-1.17) Also examined race-concordant associations (e.g., sentiment toward Hispanics and outcomes among Hispanic birthing persons, so not effect modification per-se, but results showed subgroup effects for Black and Middle Eastern birthing persons) Also, for Black birthing persons (vs full sample) the associations between negative Twitter sentiment toward Black people and birth outcomes became stronger over time (2015<2016<2017)
Hswen et al., 2020	NA	NA	NA	NA	Time series lag (autoregressive distributed) regression model	LR lag = 0.31; p = 0.022 *Interpretation:* Negative tweets mentioning Mexicans	NA	NA	NA

						and Hispanics predicted daily worry with significant lag time of one week.			

Abbreviations:
Data sources:
GSS: General Social Survey
NDI: National Death Index
IAT: Implicit Association Test
API: Application Program Interface
BRFSS: Behavioral Risk Factor Surveillance Survey
PSID: Panel Study of Income Dynamics
NCHS: National Center for Health Statistics
Geographic scales:
DMA: designated market area (media markets receiving similar media and news programming)
CBSA: Core-based statistical area (similar to metropolitan areas)
PSU: Primary sampling units (metropolitan statistical areas and nonmetropolitan counties)
Estimation:
MV: multivariable
GEE: generalized estimating equation
OLS: ordinary least squares
HLM: Hierarchical linear model
PH: proportional hazard
OR: odds ratio
PR: prevalence ratio
IR: incidence ratio
MRR: mortality rate ratio
ns: not statistically significant ($p \geq 0.05$)
Study measures:
MHI: median household income
HS: high school
NH: non-Hispanic
BMI: body mass index
PTB: preterm birth
LBW: low birthweight
VLBW: very low birthweight
CVD: cardiovascular disease
CHD: coronary heart disease
MI: myocardial infarction
Other:
E: exposure
O: outcome
C: covariates

Data source information:
GSS: The General Social Survey (GSS) is a nationally-representative sample of non-institutionalized English-speaking adults aged 18+ living in the United States conducted on a new population sample at each wave.[84,85,88,178]
BRFSS: The Behavioral Risk Factor Surveillance System (BRFSS) is a telephone-based survey (random-digit dialing of landlines and cellphones) is a telephone-based survey that focuses on chronic health conditions and health behaviors of adults across 50 states of USA and District of Columbia.[73,207]
PSID: The Panel Study on Income Dynamics (PSID) is a nationally representative, longitudinal study of households in the U.S. with interviews collected biannually by phone.[71,208]
Project Implicit: Project Implicit (PI) is a Harvard-based nonprofit research project which provides a free, online tool for assessing implicit and explicit biases toward various social groups (e.g., Blacks vs. whites, gay vs. straight people).[179] PI measures explicit biases via self-report and implicit biases via the "Implicit Association Test" (IAT).
Implicit Association Test with D-scoring algorithm: The "Implicit Association Test" (IAT) is a speeded dual-categorization task which measures the speed of keyboard associations between images of Black vs. white faces and positive (e.g., wonderful) vs. negative (e.g., disgusting) words. Faster reaction time matching positive words with white and negative words with Black faces indicates cognitive dissonance between Black people and positive emotions, which is interpreted as a pro-white implicit bias and/or anti-black Bias.[72,86,87] The IAT is scored using the D-

score measure, which ranges from -2 to +2.[209]

Explicit temperature explicit measure: Two feeling thermometer items separately ask how warm or cold participants feel toward both African Americans and European Americans (0 = very cold, 10 = very warm). Responses to the Black feeling thermometer are subtracted from responses to the white feeling thermometer, creating a score that ranges from -10 to +10 with higher values representing warmer feelings toward white people compared to Black people, interpreted as a pro-white/anti-Black explicit bias.[72,86,87]

Explicit preference measure: respondents describe how they feel toward European and African Americans using a scale that ranges from “I strongly prefer African Americans to European Americans”, to “I strongly prefer European Americans to African Americans.” Responses are on a 5-point Likert scale until 2006 and a 7-point Likert scale after 2006.[90]

Notes:

Data on Google Searches for the N-word from 2004-2007 were extracted by Seth Stephens-Davidowitz[193] using an older version of the Google Trends platform (the algorithm has since changed).

Twitter data were all geolocated with either geotag (latitude and longitude) only[89] or geotag and user-provided “place” information.[73,74]

[+] conflated sex and gender (i.e., stated they measured gender but variables were male/female (i.e., biologic sex))

Table S2. Overview of data sources used to measure area-level racial prejudice

Data source	Used in studies	Years available	Geographies available	Data access	Number aggregated	Indicator of racial prejudice
General Social Survey	Kennedy et al., 1997; Lee et al., 2015; Morey et al., 2018	1972-2018, collected every 3 years (racial attitudes questions asked beginning in 1993)	State, PSU, county, census tract	Restricted – must apply for data	Ranges from about 2,500 to 13,355 across studies	<u>Anti-Black racial prejudice:</u> Composite score based on questions: "On average, blacks have worse jobs, income, and housing. Do you think the differences are due to (a) discrimination, (b) less in-born ability to learn, (c) lack of chance for education that it takes to rise out of poverty, (d) less motivation or willpower to pull themselves out of poverty?",[84] "Do blacks tend to be unintelligent or tend to be intelligent?", and "Do blacks tend to be hard working or lazy?"[85] <u>Anti-immigrant prejudice:</u> Composite score based on questions: "Do you think the number of immigrants to America nowadays should be increased a lot, increased a little, remain the same as it is, reduced a little, or reduced a lot?" and agree or disagree with the following statements: (1) "America should take stronger measures to exclude illegal immigrants," (2) "Immigrants take jobs away from people who were born in America," (3) "Immigrants increase crime rates," and (4) "Immigrants are generally good for America's economy".[88]
Project Implicit	Leitner et al., 2016a; Leitner et al., 2016a; Orchard & Price, 2017;	2002 - present, collected continuously	State, CBSA, county	Publicly available	Ranges from about 250,000 to 1.8 million across studies	<u>Pro-White/anti-Black racial prejudice:</u> Implicit – assessed using the Implicit Association Test Explicit – assessed via self-report (temperature measure or preference measure)

	Hehman et al., 2018					All measures scored so negative values imply pro-Black/anti-White bias, positive imply pro-White/anti-Black bias, and 0 implies a neutral score.
Twitter	Nguyen et al., 2018; Huang et al., 2020; Nguyen et al., 2020; Hswen et al., 2020	2006 – present, with option for retrospective or prospective collection	Latitude + longitude available for 3-4% of public tweets, state information discernable for ~ 99% of tweets	Publicly available	1 million – 30 million	Proportion of public Tweets with latitude and longitude or other "place" information (e.g., city, state) referencing a particular racial/ethnic group that are positive, negative, or neutral. Sentiment is determined based on a combination of hand-coding, natural language processing, and machine learning.

Geographic scales:
DMA: Designated Market Area (media markets receiving similar media and news programming)
CBSA: Core-based statistical area (similar to metropolitan areas)
PSU: Primary sampling units (metropolitan statistical areas and nonmetropolitan counties)

Table S3. Strengths and limitations of data sources used to measure area-level racial prejudice

Data source	Primary strengths	Primary limitations
General Social Survey	• Nationally representative • Racial bias questions have been asked since 1993, offering greater historical context compared to the other measures* • Specificity in measurement: questions ask directly about racial attitudes* • Information on demographics of respondents (e.g., race, age, political identification, etc) is available*	• Not all questions are asked to all participants or on all survey years • Social desirability – because racial attitudes are self-reported, the GSS is subject to self-censorship or social desirability bias • Must apply for data*
Project Implicit	• Over 3-million tests have been taken since 2002 • Publicly available and free • Multiple validated tests available (e.g., racial bias, age bias, gender bias, etc.) • Can disentangle implicit vs explicit bias • Circumvent social desirability/self-censorship: IAT measures *implicit* bias through keyboard association test which does not rely on self-report • Information on demographics of test-takers (e.g., race, age, political identification, etc.) is available • Has shown high convergent validity with other measures of area-level bias*	• Project Implicit respondents are self-selected and therefore racial bias cannot be generalized to any broader population (note: some studies apply post-stratification weights on age/sex but non-representativeness on other dimensions may persist) • Repeat test-takers may regress toward the mean*
Google Trends	• Widely and regularly used by many people around the world • Circumvent social desirability/self-censorship: does not rely on self-report and search data captures *private curiosities* • Allows for real-time analysis of social attitudes* • Has been used for disease surveillance and prediction • Has shown high convergent validity with other measures of area-level bias	• Context of the search is unknown • Internet queries for the "N-Word" may not be motivated by racism • Demographics of person conducting the search are unknown* • Not possible to discern multiple searches from the same user*
Twitter	• Widely and regularly used by many people around the world • Millions of tweets are sent daily and over 90% of Twitter users make their profile and communication public • Circumvent *some* social desirability/self-censorship: does not rely on self-report and sense of anonymity may	• Geolocation data only available for small proportion of tweets where user either a) enables latitude + longitude or b) shares location of Tweets – may lead to systematic bias • Potential for residual self-censorship*:* Twitter only reflects what people were willing to express publicly

	embolden users to express views they would not display during in-person interactions • Allows for real-time analysis of social attitudes • Sentiment analysis allows researcher to characterize Tweets as positive, negative, or neutral • Has been used to characterize sentiment around a number of health topics and health outcomes	• Sentiment analysis unable to identify and process sarcasm or humor in a tweet • Demographics of person writing the Tweets are unknown

Information in this table is extracted from the 14 papers included in the systematic review. Any information that comes from content area knowledge or outside literature is indicated with an *