EXPLAINING HEALTH DISPARITIES: EVALUATING AND ADVANCING METHODOLOGY

by

Loren J. Schleiden

BS, University of Pittsburgh, 2008

Submitted to the Graduate Faculty of

the Department of Epidemiology

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

Loren J. Schleiden

It was defended on

April 3rd, 2017

and approved by

Nancy W. Glynn, PhD, Assistant Professor, Epidemiology
Graduate School of Public Health, University of Pittsburgh

Carolyn T. Thorpe, PhD, MPH, Assistant Professor, Pharmacy and Therapeutics
School of Pharmacy, University of Pittsburgh

Thesis Director: Ashley I. Naimi, PhD, Assistant Professor, Epidemiology
Graduate School of Public Health, University of Pittsburgh

Ashley I. Naimi, PhD

EXPLAINING HEALTH DISPARITIES: EVALUATING AND ADVANCING METHODOLOGY

Loren J. Schleiden, MS

University of Pittsburgh, 2017

ABSTRACT

Introduction: Health disparities are a major public health concern. Researchers often seek to explain whether these disparities are attributable to modifiable downstream risk factors that are also predictive of outcomes. Traditionally-used mediation approaches may not be appropriate to explain such disparities, due to the complex nature of socio-demographic factors such as race/ethnicity.

Methods: Aim 1: We carried out a methodological review of the literature. We searched for studies that investigated a racial/ethnic disparity in a health outcome and adjusted for a potential modifiable mediator of the disparity. Out of the 969 studies meeting inclusion criteria, a simple random sample of fifty studies were drawn to undergo thorough data abstraction. Data were abstracted to determine what mediation approaches were employed, if causal language was used to describe these estimates, and if relevant assumptions were considered in justifying this causal language. Aim 2: We also carried out an analysis of National Survey of Family Growth (NSFG) data representing 103,919,318 live singleton births to mothers in the United States. This analysis employed traditional (Difference and Product methods) and general mediation methods (Inverse Probability-Weighted Marginal Structural Models and Structural Transformation) to result in different estimates of how payment method for delivery explained the racial/ethnic disparity in low birth weight.

Results: All studies included in the target review employed traditional mediation approaches, the majority used language that could be interpreted as causal (86%), and none of the studies explicitly addressed assumptions or conditions for interpreting estimates causally. In the analysis of NSFG data, for every 100 births, there were an additional 5.27 low birth weight infants born to non-Hispanic Black mothers, compared to births to non-Hispanic White mothers. Estimates of the counterfactual disparity measure ranged from 1.62 using the product method to 4.66 using the structural transformation method.

Conclusion: Traditional mediation approaches are commonly used to explain racial/ethnic disparities. The use of such traditional approaches to explain disparities along the lines of complex socio-demographic variables may be inappropriate due to underlying assumptions that are likely broken. Considerations of these underlying assumptions and use of appropriate mediation approaches are crucial to understanding the important public health issue of health disparities.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank the faculty and staff at the Graduate School of Public Health for their support throughout my student career. I would also like to acknowledge and thank Drs. Naimi, Glynn, and Thorpe for their feedback and my effort throughout my thesis.

1.0 INTRODUCTION

The Problem of Health Disparities

Health disparities are among the most salient public health concerns in the United States. Cancer, HIV/AIDS, homicide, cardiovascular, reproductive, and pediatric health outcomes are all strongly patterned along socio-demographic lines, with disadvantaged populations carrying a disproportionate disease burden.[1] Such socio-demographic factors that distinguish advantaged from disadvantaged populations include (but are not limited to) sex/gender, sexual orientation, education, income, occupation, and race/ethnicity. These socio-demographic factors have been deemed "upstream" because they are thought to influence more proximal ("downstream") determinants of health that include health behaviors, stress, and access to care, which subsequently pattern health outcomes.[2]

Despite many of these health disparities being long-demonstrated, intervening to reduce such disparities is difficult. Health disparities are multifaceted and can thus be impacted by the interplay of a multitude of social determinants. Contemporary social determinants are seen as the result of a long history of social processes and institutional mechanisms. Many of these disparities are nuanced and may not be appropriately captured through the convenience measures that are often used to report them. Due to their complex nature, it can be impossible to fully understand health disparities, much less to implement interventions that fully address them. As a result, many health disparities that have been observed for decades persist, and there have

recently been calls for further research and policy changes.[3] Healthy People 2020 has made it an overarching goal to "achieve health equity, eliminate disparities, and improve the health of all groups,"[4] and reducing disparities is a core objective of the Patient Protection and Affordable Care Act (ACA; Pub L No. 11-148).

"Explaining" Health Disparities

Addressing the nature of health disparities is crucial. To that end, there is a large, long-standing, and continually growing body of research devoted to "explaining" why health disparities exist. This work often seeks to account for the extent to which an overall statistical association between a health disparity marker and health outcome is attributable to other factors. For example, a 2007 study of data from the National Health and Nutrition Examination Surveys attempted to "explain" excess hypertension risk among African Americans and US-born Hispanics (compared to Whites) by controlling for measures of socioeconomic status (education, household income, and poverty income ratio), health behavior (lack of exercise, current smoking, and poor diet), and access to care (current health insurance availability).[5] In this study, researchers sought to evaluate how much the racial/ethnic disparity in hypertension risk was attributable to the corresponding racial differences in these downstream factors (socioeconomic status, health behavior, and access to care).

Due to the complex nature of these socio-demographic factors, typical markers of a given health disparity, such as race/ethnicity, cannot be interpreted as corresponding to some well-defined intervention. However, social, clinical, and/or public health policy changes may be able to target modifiable "downstream" factors that act as proximal risk factors that play a role in transmitting the association between the health disparity marker (exposure) and outcome. For example, a systematic review of interventions targeting racial/ethnic disparities in diabetes risk

presented patient interventions targeting potentially modifiable downstream risk factors such as dietary habits, physical activity, and diabetes self-management activities.[6] The broader goal of this type of research is to influence policy change, resulting in interventions that can lead to a reduction in health disparities. However, it is important that these policy changes are informed by best research practices. In order to develop useful empirical evidence, researchers must use analytic methods that properly answer pertinent clinical and public health questions.

Mediation Analysis

Questions pertaining to the extent to which a health disparity is "explained" by a modifiable downstream risk factor are fundamentally questions about mediation. Mediation analysis refers to a set of analytic techniques that enable researchers to evaluate the processes or mechanisms by which an exposure is associated with an outcome. A common approach to explaining health disparities has sought to identify modifiable mediators, which, if altered, will lead to a reduction in health disparities. This approach is based on the presumption that the information obtained will valuably contribute evidence on whether an available intervention should be used to reduce health disparities and improve overall health. This presumption requires a logical correspondence between the research question, analytic method, and the potential clinical or public health interventions. The common objective of all such analyses is thus to determine whether an intervention targeting the mediator in question will make a viable and worthwhile impact.

Despite this common objective, a multitude of disparate analytic approaches have been used to quantify the extent to which mediators explain health disparities. These approaches involve an attempt to either: (1) decompose the effects into direct and indirect components; or (2) quantify the disparity that would remain if the mediator were set at a single value within a

population. The direct exposure effect represents the effect of the exposure on the outcome that does not operate through a mediator [e.g., the effect of race/ethnicity (exposure) on hypertension (outcome) risk independent of income (mediator)]. The indirect exposure effect quantifies the effect of the exposure on the outcome that operates through the exposure's impact on a mediator [e.g. the effect of race/ethnicity (exposure) on hypertension risk (outcome) through its impact on income (mediator)]. These are distinct from the effect that would remain if the mediator were set to a specific value uniformly in the population.

The Difference Method

Although a large number of analytic techniques are available to conduct mediation analysis, researchers often rely on a select few. One technique known as the "difference method" has long been used as a template to conduct such analyses. With the difference method, the effect

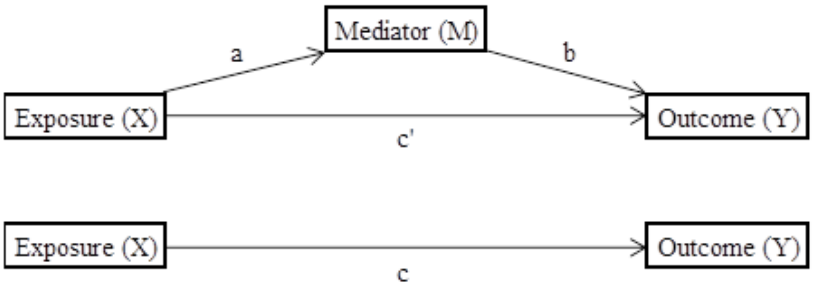

Figure 1. Difference Method Mediation Model

of an exposure on an outcome without adjusting for a mediator (illustrated as c in Figure 1) is compared to the exposure's effect after adjusting for a mediator (illustrated as c' in Figure 1). Regarding two linear regression models of the outcome (Y) that include the exposure (X), with one model not adjusting for the proposed mediator(s) (M) and another model adjusting for M:

1. $Y = \beta 2 + cX + \varepsilon_1$
2. $Y = \beta 3 + c'X + bM + \varepsilon_2$

According to the difference method, the difference in the effect of the exposure on the outcome before and after adjusting for the mediator(s) (coefficients c and c' in the equations above) is considered to be the indirect effect (i.e., effect of the exposure that operates through the mediator), while the effect that remains after adjusting for the mediator is considered the direct effect (i.e., the effect of the exposure that does not operate through the mediator). In practice, the difference method will only provide valid results under linear regression models and under stringent no confounding and no interaction conditions (discussed further later). However, non-linear regression methods (e.g., logistic, Cox proportional hazards) are often used, and the conditions are often ignored.

The Product Method

Similar to the difference method is the "product method" initially proposed by Baron and Kenny.[7] According to Google Scholar, Baron and Kenny's 1986 methodological paper on the moderator-mediator variable distinction has been cited 62,558 times, which speaks to how widely the product method has been used in the literature. In Figure 1, the product method is depicted as decomposing effects into the direct effect (c) and indirect effect (product of coefficients a and b).

According to the product method, mediation can be evaluated using three regression models (equations shown below, where a, b, c, and c' correspond to the Figure 1 above): 1) a model of the exposure predicting the outcome; 2) a model of the exposure predicting the mediator; and 3) a model of the exposure predicting the outcome while adjusting for the mediator:

1. $Y = \beta_1 + cX + \varepsilon_1$
2. $M = \beta_2 + aX + \varepsilon_2$

3. $Y = \beta_3 + c'X + bM + \varepsilon_3$

According to the product method, significant associations need to be confirmed between the exposure and outcome in the unadjusted model predicting the outcome (c), between the exposure and mediator in the model predicting the mediator (a), and between the mediator and outcome in the model including the exposure and mediator predicting the outcome (b). A decrease in the association between the exposure and outcome going from the model that does not include the mediator (c) to the model adjusting for the mediator (c') is taken as evidence of mediation.

Problems of the Difference and Product Methods

While the difference method and product method seem intuitive, closer scrutiny suggests that they cannot be used to analyze health disparities. A health disparity is defined as a comparison (e.g., difference, ratio) of the measure of occurrence between two groups identified by some characteristic. For example, black men are at 1.23 times the risk of heart disease of white men in the United States.[8] Such a quantity can be estimated using standard unadjusted regression without issue. However, when researchers seek to "explain" health disparities, the target quantity of interest is often framed as an answer to questions such as: what would the magnitude of the disparity be if some downstream risk factor were held constant at a specified value?

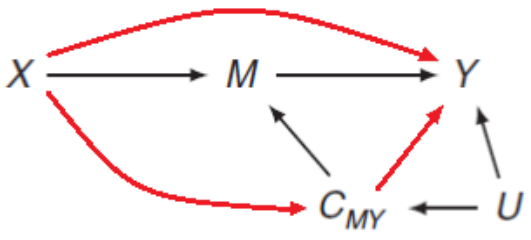

Figure 2. Mediation Diagram with Measured Confounder of the Mediator-Outcome (C_{my})

This question can be answered by quantifying the red arrows in Figure 2. This figure shows the relations between the sociodemographic characteristic of interest (X), the downstream risk factor (M) and outcome (Y) under study, and confounders of the relation between the downstream risk factor and the outcome (C_{MY}). In this figure, the disparity that would remain if the downstream risk factor were set to a specific level is represented by all (red) arrows emanating from X that are not transmitted though M (i.e., $X \rightarrow Y$ *and* $X \rightarrow C_{MY} \rightarrow Y$).

To quantify the racial/ethnic disparity that would remain if M were intervened upon using standard approaches (difference or product methods), several strong assumptions are required. These include:[9]

1. No uncontrolled mediator-outcome confounding
2. No exposure-mediator interaction on the scale of interest
3. No mediator-outcome confounders affected by the exposure

The single-most important complication that arises when seeking to quantify the extent to which a health disparity is explained by a downstream risk factor is the violation of assumption 3, represented by the $X \rightarrow C_{MY}$ relation in Figure 2. Socio-demographic characteristics denoted X in the figure (e.g., race/ethnicity) are deemed "upstream" precisely because they are associated with a wide range of risk factors. These risk factors include the primary intervention variable of interest M, as well as any variables that confound the M→Y relation (C_{MY}). Because these latter variables are confounders, they must be adjusted for in any analyses seeking to quantify the disparity that would remain if M were set to a specific level in the population. However, adjusting for them using standard techniques such as the difference or product methods will also block part of the association of interest (i.e., $X \rightarrow C_{MY} \rightarrow Y$). This issue represents a fundamental analytic challenge in explaining health disparities.

Additionally, controlling for this mediator-outcome confounder via standard regression methods in the presence of further unmeasured confounding between this confounder and the outcome would induce a collider bias, opening a pathway from exposure to the outcome ($X \rightarrow C_{MY} \leftarrow U \rightarrow Y$ in the figure above). Collider bias can occur when conditioning on a variable that is an outcome of both the exposure and the outcome, and it has the opposite effect of conditioning on a non-collider: conditioning on a non-collider blocks the flow of association along a path, while conditioning on a collider opens the flow of association.[10] Finally, the presence of an exposure-mediator interaction, thus violating assumption 3, could render causal inferences made through the difference method invalid. Exposure-mediator interaction makes it impossible to measure a single direct effect, because the direct effects vary by different levels of the mediator.

Importantly, the violations of each of these assumptions is highly likely in a health disparities setting due to the nature of sociodemographic variables, and the structural relations that exist, as demonstrated in Figure 1. Consequently, despite their frequent use, the difference and product methods are inappropriate for answering key questions of interest health disparities research.

An additional complication arises with use of the difference and product methods. When using an association measure that is non-collapsible, such as an odds ratio or hazard ratio, the unadjusted odds ratio or hazard ratio cannot be compared to an odds ratio or hazard ratio that has been adjusted for a mediator.[11] This may be more problematic when considering a dichotomous outcome that is common, as the impact of non-collapsibility is more severe. [12] There do exist circumstances in which the difference method and the product method yield valid causal effects.

But the assumptions required are strong and unverifiable. It is thus not possible for investigators and readers of the health disparities literature to know when these assumptions have been met.

General Mediation Approaches

Other approaches to mediation analysis have recently been developed, many of which overcome the limitations of the difference method and product methods. Inverse probability-weighted marginal structural models (IPW MSMs) have been proposed as a mediation method that can be used to properly adjust for mediator-outcome confounders associated with the exposure.[13] By generating inverse probability weights from models of the mediator and fitting a weighted regression model of the outcome that includes the exposure, mediator, and their interaction, mediator-outcome confounders can be properly accounted for. Structural transformation accounts for confounding through first modeling the mediator. This model can provide an unbiased estimate of the mediation effect. From this model, parameters of the mediator and exposure-mediator interaction term are subtracted from the outcome to create a transformed outcome. In effect, this act of transforming the outcomes removes the effect of the mediator. The outcome is then modeled against the exposure to estimate the effect of the exposure on the outcome, with the effect of the mediator removed. G-estimation of a structural nested mean model[14] and targeted minimum loss-based estimation (TMLE)[15] each model both the mediator the outcome model, and combine these models to yield doubly robust estimators, which are consistent if either the mediator or outcome (but not necessarily both) model is correctly specified.[16-18]

Considering Mediation Techniques in Analysis of Racial/Ethnic Disparities

The distinctions between the difference or product methods and more general approaches can have important consequences on research findings. Indeed, using these various mediation

analysis techniques can produce widely varying results. As shown by Naimi et al, using the difference method and generalized product method versus more advanced methods (IPW MSMs, structural transformation, g-estimation, and TMLE) resulted in drastically different estimates of the counterfactual disparity measure. Estimates of the proportion of the disparity explained ranged from 126% using the difference method to 8% using structural transformation.[19]

The overarching goal of many health disparities investigators is to influence policy changes at the health system, government, and societal levels to reduce disparities and improve health outcomes. When spurious results are presented as causal effects, even though the criteria for making a causal statement have not been met or considered, the influence on policy change could lead to ineffective interventions, waste of resources, and further confusion on the already complex nature of health disparities.

Racial Disparities in the United States

To better focus the remainder of the current work, the state of the research on racial disparities in health in the United States will be considered. Race/ethnicity is a non-modifiable socio-demographic factor along which a multitude of health disparities have been patterned. Various studies have sought to explain racial/ethnic disparities using proposed mediators.[20-22] The crucial roles that race/ethnicity play in health outcomes demand that it continue to be addressed through research and policy change, but researchers often fail to consider the approaches and analytic methods that are appropriate in estimating the role of mediators in these disparities. Racial/ethnic health disparities in the United States are pervasive and severe, with national disparities by race/ethnicity in infant mortality,[23] heart disease,[24] cancer,[20,25,26] HIV,[27] and life expectancy.[28] As such, race/ethnicity is one of the most salient socio-demographic factors in the United States, and it will be the focus of this thesis.

Objectives

The objective of this thesis is two-fold: Aim 1of the thesis is to conduct a targeted review of the literature in order to describe the analytic strategies used to quantify the extent to which racial/ethnic health disparities are explained by more proximal risk factors. This review will focus on the following research questions: 1) what analytic techniques are being used to quantify the effect of proximal risk factors in racial/ethnic health disparities?; 2) what effect estimates are being reported, and were techniques used appropriate for reporting such effects?; and 3) were causal explanations implied in the language used to describe different effect estimates, and was this use justified? No such targeted review of the literature has been published.

The objective of Aim 2 is to conduct an analysis using traditional and general mediation approaches to quantify the same target quantity. This will demonstrate how applying different mediation approaches can result in a range of estimates of the mediation effect when applied to a question of racial/ethnic disparity. The National Survey of Family Growth (NSFG) will be used to answer the research question "how much does payment method explain the racial disparity in low birth weight births?" This research question will be answered using traditional (difference method and product method) and general (IPW MSMs and structural transformation) mediation approaches to determine estimates of the counterfactual disparity measure and proportion of the disparity explained. The counterfactual disparity measure (CDM) can be interpreted as the magnitude of the disparity if, contrary to fact, the mediator could be set to a certain level. Proportion explained is the estimate of the proportion of the total effect of the exposure on the outcome that is explained by a mediator.

2.0 METHODS

Aim 1 - Target Review

The targeted review was conducted in accordance with a prospective protocol designed for this study. This protocol was in part guided by a Public Health Informationist at the University of Pittsburgh Health Sciences Library System (HSLS). Inclusion criteria were established by authors LS and AN based on preliminary searches. Five exemplar studies meeting inclusion criteria were determined and checked for in all searches. The abstracts of final results were inspected by one reviewer to determine evidence of inclusion criteria, with full texts being consulted for any results that could not be determined through abstract review alone. Inclusion criteria were verified by a second reviewer for 10% of final results. Pertinent information was abstracted from full texts for a simple random sample of 50 articles meeting inclusion criteria.

Inclusion Criteria

Studies were considered to have met inclusion criteria if they: 1) investigated at least one racial and/or ethnic disparity in a health condition, service, treatment, or outcome; 2) adjusted for a variable or group of variables interpreted as modifiable mediator(s) of the measure of race/ethnicity and outcome association (i.e, race/ethnicity is associated with a mediator variable(s), which in turn impacts the outcome, thus “explaining” at least a proportion of the association between race/ethnicity and outcome); and 3) were published between 01/01/2000 and

05/18/2016. Reviews or meta-analyses without original data, case series or reports, comments, letters to the editor, and editorials were excluded from the review.

Identification of Studies

A search was conducted by the investigators in PubMed combining general and key word terms pertaining to mediation analyses, race/ethnicity, and disparities, informed in part by Medical Subject Heading (MeSH) terms. Terms for mediation analysis included *mediation analy*, mediator analy*, controlled direct, natural direct, natural indirect, natural effect, direct effect, proportion explain*,* and *explain*.* Terms for race/ethnicity included *race, racial,* and *ethnic*.* Terms for disparities included *disparit*, social determinant*,* and *social condition*.* These search terms identified a total of 1,913 results through PubMed.

Verification of Inclusion Criteria

Bibliographic details and abstracts from resulting citations were downloaded. One reviewer screened all citations based on the abstract to identify studies meeting inclusion criteria. The reviewer retrieved full texts for any studies whose inclusion could not be determined through information provided in the abstract.

Of the 1,913 results, 969 were found to meet inclusion criteria. A simple random sample of 192 articles (10%) identified through search terms were reviewed by a second reviewer. Percentage agreement (90.6%) and Kappa statistics (K = 0.803, 95% CI: 0.717-0.89) were calculated to reflect the degree of agreement, and the strength of agreement was considered to be very good. Any discrepancies between the two reviewers were resolved through discussion, and articles meeting inclusion criteria were eligible for the random sample of publications from which data would be abstracted.

Data Abstraction

Full-text articles for a simple random sample of 50 out of 969 publications meeting inclusion criteria were retrieved. Upon further detailed inspection, any publications that were found not to meet inclusion criteria were replaced with another article selected via simple random selection. This was the case for two out of the 50 articles. Two investigators worked to abstract pertinent study data into a single data collection spreadsheet created for the purposes of this study. First author, year, title, race/ethnicity categories, mediators, outcome, other covariates included in analyses, mediation analysis methods cited or used, mediation analysis results (including: total effects, direct effects, indirect effects, and proportion of disparity eliminated), and evidence of causal statements were sought from each publication. Mediation analysis methods were grouped into three categories: 1) Difference method; 2) Product method; 3) General mediation (including but not limited to G-estimation of structural nested models, targeted minimum loss-based estimation (TMLE), and inverse probability weighted marginal structural models (IPW MSMs).

Aim 2 - Mediation Analyses

Aim 2 consists of applying different mediation approaches to the same research question while using the same data, in order to determine the variation in results when using different mediation approaches. The research question across analyses is "How much does delivery payment method account for the racial/ethnic disparity in low birth weight between infants born to non-Hispanic white and non-Hispanic black mothers?" This research question provides an example of a well-documented racial disparity that a researcher might want to explain using one or more potentially modifiable mediators.[29-32] Infants born at low birth weight (less than 2500 grams or 5.5 pounds) are at increased risk of developmental delays, long-term complications, and

infant death. Low birth weight infants are often also born preterm (before 37 completed weeks of gestation), but they could have also experienced fetal growth restriction. Preterm birth and fetal growth restriction can be associated with infection, maternal medical conditions, exposure to tobacco smoke, and inadequate maternal weight gain during pregnancy. There exists a national disparity in low birth weight, as 13.2% of infants born to non-Hispanic Black mothers and 7.0% of infants born to non-Hispanic White mothers are low birth weight.[32] Previous studies attempting to explain the racial disparity in low birth weight have controlled for variables such as maternal health,[33,34] stress,[29,35] and socioeconomic variables such as using government insurance to pay for birth.[35] As a potentially modifiable "downstream" factor, payment method for delivery is associated with race/ethnicity and in turn has a potential association with the outcome of low birth weight.

Analyses were conducted on a sample from the 2006-2010 and 2011-2013 waves of the National Survey of Family Growth (NSFG), a nationally representative survey of individuals age 15-44 years in the United States. The NSFG waves 2006-2010 and 2011-2013 contain information on 30,035 pregnancies. As illustrated in Figure 3, analyses were limited to singleton livebirths, resulting in a sample of 20,666. For the purposes of these analyses, individuals of Hispanic and non-Hispanic Other Race/Ethnicity were also excluded, resulting in a final cohort of 13,611 singleton live births. With survey weights applied, these 13,611 births represent 103,919,318 births to non-Hispanic White or non-Hispanic Black mothers in the United States. Race/ethnicity of mothers (non-Hispanic white (NHW) or non-Hispanic black (NHB)) was self-reported during NSFG interviews, and will be used as the "upstream" determinant of health in these analyses. The outcome was low birth weight (LBW), categorized into a binary variable of

2,499 grams or less versus 2,500 grams or more, as according to the World Health Organization definition. The proposed mediator, payment method for delivery, was categorized as a binary

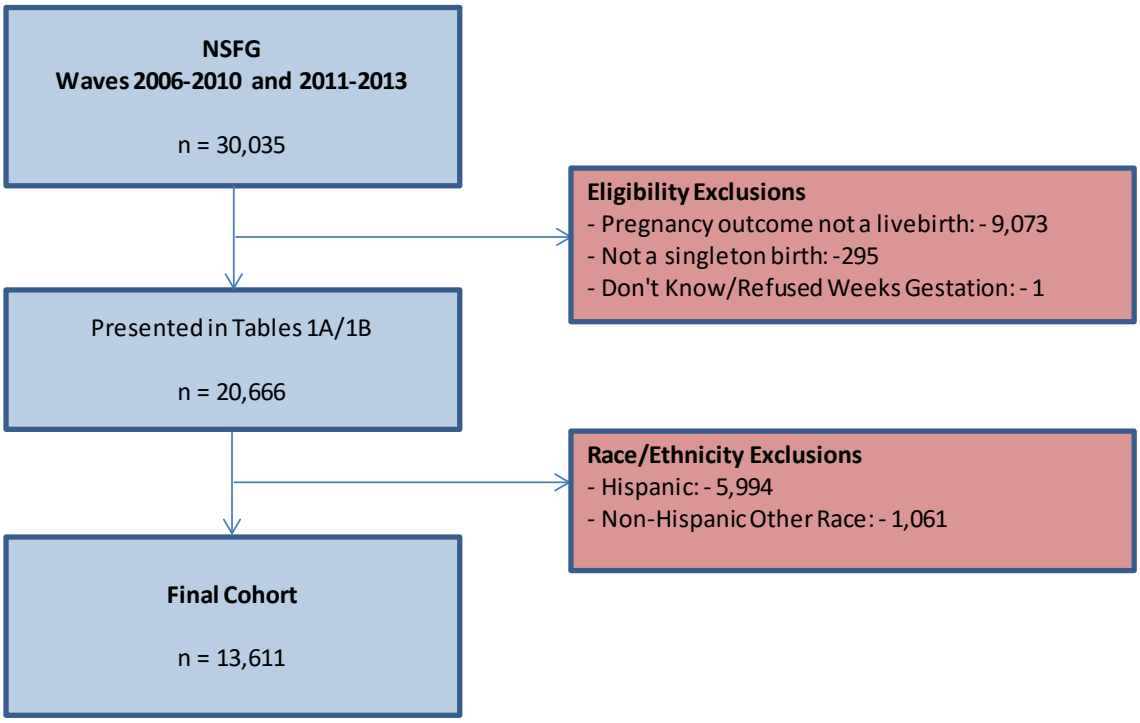

Figure 3. National Survey of Family Growth Cohort Flowchart

variable categorized as private insurance/own income versus Medicaid/Government Assistance (referent value). Other covariates included in the model to represent confounders of the mediator and outcome were maternal age at birth (under 20 years, 20-24 years, 25-29 years, and 30-44 years old (referent value), maternal education (<12th grade vs HS degree or more (referent value)), maternal marital status (divorced/separated/cohabitating/never married versus married (referent value)), prenatal care (received before 20th week of gestation (referent value) versus not received/received after 20 weeks of gestation), gestational length (before 36 weeks gestation versus at least 36 weeks gestation (referent value), wantedness of pregnancy (unwanted versus

right time/overdue/mistimed/don't know (referent value)), and infant born in the United States (yes (referent value) vs no).

Mediation approaches demonstrated through example analyses included the difference method, product method, and general methods, in this case consisting of IPW MSMs and structural transformation. For the outcome of LBW, estimates of the NHW versus NHB disparity in fully adjusted models using each of the mediation approaches were compared to the risk difference in a model containing only race/ethnicity to estimate the counterfactual disparity measure (CDM) and disparity "proportion explained." The CDM is defined as the unadjusted racial/ethnic disparity in LBW that was observed minus the LBW disparity that would exist if all mothers had paid for their delivery through own income, insurance, or a combination of own income/insurance. As a potential outcome, this latter disparity is impossible to observe. However, under given assumptions, it is possible to quantify on average using various mediation approaches. The proportion explained is conceptualized as the proportion of the unadjusted disparity that would be attenuated (or possibly exacerbated) if the mediator could be set to a certain level. For these example analyses, the proportion explained is calculated by subtracting the estimation of the race/ethnicity CDM for each mediation approach by the estimation of the unadjusted race/ethnicity CDM, dividing by the unadjusted race/ethnicity CDM, and multiplying by 100 to estimate a percent.

$$\text{Proportion Explained} = (CDM_{\text{unadjusted}} - CDM_{\text{mediation approach}})/ CDM_{\text{unadjusted}} \times 100$$

3.0 RESULTS

Aim 1 - Targeted Review

Although articles published between 01/01/2000 and 05/18/2016 were considered eligible, 540 out of the 969 eligible articles identified (55.7%) and 31 of the simple random sample of 50 (62.0%) articles were published in 2010 or later. The majority of these 50 studies were conducted in the United States (43), while two were conducted in New Zealand and one each in Australia, Brazil, Iceland, the Netherlands, and South Africa. All studies used an observational cohort design, with 37 being retrospective and 13 having prospective designs. Studies using retrospective cohort designs employed data sources such as Medicare/Medicaid claims, National Health Interview Survey, National Health and Nutrition Examination Survey, and records from regional health-systems.

Studies compared a variety of race/ethnicity groups, and there were many inconsistent coding schemes used to categorize race/ethnicity. The most common race/ethnicity categorization was simply white and black, as 20 studies included only these two groups for comparison. Multiple studies conducted in the United States also included Hispanic and/or Asian racial/ethnic groups, and one study each also used American Indian, Indian/Pakistani, and "Other" racial/ethnic groups, respectively. While some studies specified "non-Hispanic" white or black, others did not make this distinction. Studies conducted outside of the United States typically used very different racial/ethnic categories as studies conducted in the United States.

The two studies conducted in New Zealand included Maori, New Zealand European, and Asian race/ethnicities. The study conducted in South Africa categorized race/ethnicity as white, black, Indian, and "colored," while the study conducted in Brazil compared white, black, and "brown" racial/ethnic groups. The Icelandic study compared non-mixed Icelandic, non-mixed Polish, mixed Polish, non-mixed Asian, and mixed Asian racial/ethnic groups.

Mediation Approaches Used

Of the simple random sample of 50 articles, 49 used the difference method to estimate the effect of the proposed mediator(s) on the racial/ethnic disparity in a health outcome. The one study that was an exception used the generalized product method, although this was not explicitly stated, nor were Baron-Kenny cited. No studies used mediation approaches categorized as general methods. Of studies that used the difference method, 21 used multivariate logistic regression, 15 used multivariate linear regression, and 9 used Cox regression/Kaplan-Meier survival analysis. One study each employed structural equation models, multivariate path models, multinomial logistic regression, and chi-square tests. The one study that used the generalized product method employed multivariate logistic regression.

Underlying Assumptions Required

None of the studies identified explicitly confirmed or acknowledged any of the assumptions needed in order to estimate causal effects using the difference and/or product methods (i.e., no uncontrolled mediator-outcome confounding; no mediator-outcome confounders affected by the exposure; no exposure-mediator interaction on the scale of interest). One study by Myakovsky (2012) makes the statement "All data were examined for statistical assumptions," but does not go into detail about what these assumptions were, nor are the assumptions for mediation analyses specifically mentioned.[36] Articles by Do (2012) and Beck

(2014) had a greater focus on methodology of SES as a factor to explain racial/disparities and on decomposing racial/ethnic disparities, respectively, but these articles do not specifically address the assumptions necessary in order to estimate causal effects using the difference or product methods.[37,38] Despite not addressing these assumptions, the majority of identified studies used language that could be interpreted as causal (43/50), using words such as "explain," "attenuate," "contribute," and "account for" when describing the effect of the mediator on the racial/ethnic disparity. For example, article by Runarsdottir states that "ethnic differences in life-satisfaction and distress were almost fully accounted for when sociodemographic background and social support were controlled."[39] Studies that did not use causal language used words such as "associated" or used phrases such as "may have an influence," "likely involved," or "may contribute" when describing the effect of the mediator on the racial/ethnic disparity.

Proportion Explained

Five studies reported a version "proportion explained" of the disparity. This estimate connotes the amount the disparity would change if the mediator(s) were set to a certain level. The typical method to calculate this proportion explained involved subtracting the estimate of the race/ethnicity effect in the model controlling for the proposed mediator(s) from the estimate of the race/ethnicity effect in the crude model, and then dividing this difference by the crude model estimate of the race/ethnicity effect. The formula used to calculate the proportion explained varied across studies, as some studies compared Odds Ratios of the race/ethnicity effect while others compared coefficients. One study by de Hoog did not disclose the formula used to calculate the average proportion explained across race/ethnicity groups.[40] The article by Osborne treats multiple proposed mediators as each explaining a percentage of the racial ethnic disparity, even though they were added in a hierarchical fashion.[41] "We estimate that 29% of the observed

disparities in mortality is caused by patient comorbidities, 6% from use of endovascular repairs, 26% due to socioeconomic factors, and 25% because black patients receive care in lower-quality hospitals." Since they were added hierarchically, estimates of the proportion explained by patient comorbidities could change in models that also control for the other proposed mediators. Authors later state that “still unexplained are 14% of the difference in mortality between black and non-black patients.” Treating proportion explained in this way could be problematic, as one might infer that a disparity has been completely accounted for when the proportion explained reaches 100%. However, the proportion explained is not a true proportion, as calculating the proportion explained treats the total effect as a constant. In actuality, the total effect can drastically change due to sampling variability. This can lead to very different estimates of proportion explained. Proportion explained can also cause problems for interpretation when estimating a value of less than 0% or more than 100%.[42]

All of the studies included in this targeted review were attempting to explain a racial/ethnic disparity using modifiable mediating variables. Every study used traditional mediation approaches, but not one study used a mediation approach that appropriately accounts for the required underlying assumptions. Most notably, no study accounts for the probable association of the exposure (race/ethnicity) with a confounder of the mediator-outcome association. This assumption was not considered, so it is likely that the traditional methods were not appropriate, and thus estimates of the mediation effects are not accurate. It is important to consider the degree to which these results could be flawed by understanding the range of estimates that could be found through applying various mediation approaches to the same data in order to answer the same research question.

Aim 2 - Mediation Analyses

Survey-weighted characteristics of the NSFG sample are presented in Table 1. With survey weights applied, the sample represents 103,919,318 births. In this sample restricted to NHW and NHB mothers, 77.1% of births were to NHW mothers and 22.9% were to NHB mothers. Over one third of mothers paid for their delivery using only Medicaid/Government Assistance (36.7%). Most women were in their twenties when they gave birth, but 23.6% of births were to mothers under 20 years old and another 17.1% of births were to mothers over 30 years old. Nearly one sixth of mothers (14.0%) had below a 12^{th} grade education and 47.6% were not married at the time of birth. Nearly all mothers had received prenatal care before the 20^{th} week of gestation (95.4%), had a gestational length of at least 36 weeks (91.8%), and did not believe their pregnancy was unwanted (85.1%). Nearly all infants were born in the United States (95.4%).

Overall, 7.7% of births were LBW, including 6.5% of births to NHW mothers 11.8% of births to NHB mothers. The survey-weighted unadjusted risk ratio was 1.81 (95% Confidence Interval (CI): 1.53, 2.14), with infants born to NHB mothers being at almost double the risk of LBW as infants born to NHW mothers. Births that were paid for by Medicaid/Government Assistance were associated with an unadjusted LBW risk difference of 0.04 and a risk ratio of 1.68 (95% CI: 1.52, 2.01), connoting a greater risk than births paid for with insurance, own income, or a combination of income/insurance. Younger maternal age at birth, less than 12^{th} grade education, being not married, not having prenatal care before 20^{th} week of gestation, gestational length <36 weeks, unwanted pregnancy, and births outside of the United States were all associated with greater risk of LBW events in unadjusted bivariate analyses.

The estimated unadjusted LBW risk difference among births to NHB and NHW mothers and CDMs using various mediation analyses are presented in Table 3. Applying survey weights to a simple linear regression model containing only race/ethnicity as a predictor resulted in an estimated 5.27 (95% CI: 3.65, 6.89) additional deaths for every 100 births to NHB mothers compared with births to NHW mothers. Using the difference method to account for the proposed mediator of payment method suggested that the disparity would be reduced 58.8% to an additional 2.20 (95% CI: 0.64, 3.70) deaths for every 100 births to NHB mothers if all mothers had a payment method of private insurance/own income (vs Government Assistance/Medicaid). The generalized product method estimated a proportion explained of 69.2% (CDM 1.62, 95% CI: -0.44, 3.68). The general methods had lower estimates of proportion explained, as using a IPW MSMs mediation approach suggested a proportion explained of 31.6% (CDM 3.61, 95% CI: 1.46, 5.75), and using a Structural Transformation mediation approach suggested a proportion explained of 11.5% (CDM 4.66, 95% CI: 3.05, 6.28).

4.0 DISCUSSION

The ultimate goal of health disparities research is not simply to measure a disparity, but to deepen understanding about the nature of a disparity so that some action can be taken to reduce it.[43] While it may be impossible to intervene on complex upstream socio-demographic factors such as race/ethnicity, mediation analysis approaches enable researchers to demonstrate the effects of associated modifiable intermediate variables that in turn have an effect on a health outcome of interest. In theory, a well-designed intervention targeting these mediators would be able to reduce the disparity based on the socio-demographic factor of interest. Though these concepts may seem simple, applying appropriate mediation analysis techniques can become complicated when dealing with convoluted socio-demographic factors such as race/ethnicity.

It is a well-established principle in the philosophy of science that no hypothesis can be tested in isolation, since any one hypothesis relies on the assumption that multiple other hypotheses are correct.[44] Epidemiology is no exception, and many of the analytic techniques practiced in epidemiology were built upon the assumption that a multitude of underlying hypotheses are also correct. For example, an underlying assumption of linear regression is that there is a linear relationship between independent variable(s) and the dependent variable. When this assumption is violated, estimates resulting from linear regression are no longer appropriate for interpretation. In order to appropriately approach the complexity in applying mediation analysis techniques to disparities along socio-demographic factors, there needs to be careful

consideration of the underlying assumptions required in order to test a hypothesis of interest. Approaching mediation analyses from a causal inference framework[45] can allow models to be depicted graphically using directed acyclic graphs (DAGs) and scrutinized for various causal pathways implied by the model.[46] Through this causal inference lens, traditional mediation methods such as the difference method and product method are shown to require much more stringent assumptions that can be difficult to meet and impossible to verify when dealing with a complicated socio-demographic exposures.

Despite these assumptions that are challenging to presume in order to appropriately employ traditional mediation approaches, every study included in the targeted review used traditional mediation approaches. Furthermore, none of these studies explicitly discussed or addressed any of these assumptions, even though the vast majority used language that could be interpreted as implying a causal effect. While using language such as “explained” and “accounted for” may seem innocuous, this sends a message to readers that these effects are causal, and thus a real-life change in these factors would result in the estimated change in the disparity. If these assumptions are in fact violated, these estimates could be drastically different. The example mediation analyses in this thesis have demonstrated important differences in counterfactual disparity measure according to the applied mediation analyses approaches applied. This is due to how these traditional and general methods account for confounding in different ways, as discussed on page 9.

Employing the different mediation approaches resulted in a wide range of estimates of CDM and proportion explained, with the proportion explained ranging from 69.2% using the product method to 11.5% using structural transformation. This range of estimates is attributed to differences in how the traditional and general approaches handle confounding. In the case of

racial/ethnic disparities, perhaps the most difficult assumption to hold is that of there being no confounders of the mediator and outcome association that are in turn associated with the exposure. This is due to the nearly universal impact of race/ethnicity on downstream factors and outcomes. Since it is likely that there is an association between race/ethnicity and any potential confounders of a proposed mediator and outcome, it is likely going to be necessary to use mediation approaches that do not simply control for these confounders.

In effect, IPW-MSMs account for this confounding by removing the association between the exposure and mediator and removing the association between confounders of the mediator-outcome association and the mediator. Structural transformation involves transformation of the outcome in order to subtract the effect of the mediator from the outcome. This transformed outcome is then modeled against the exposure.[47] An in-depth explanation of how these general mediation approaches account for confounders of the mediator-outcome association are described elsewhere.[9,19,48] These approaches involve be separate assumptions required in order to interpret estimates as causal effects. For instance, IPW-MSMs require the assumption that the mediator model is correctly specified as a function of all confounders of the mediator-outcome association, and structural transformation requires that the outcome model is correctly specified. However, these assumptions may be considerably easier to believe compared to the assumption that race/ethnicity would not have any effect on any confounder of the mediator-outcome association.

As described by Beyea and Greenland (1999), analyses should test the robustness of estimates by repeating them for a broad range of underlying biologic models.[49] Along the lines of this logic, it can be important to use multiple mediation approaches in order to account for a variety of underlying causal models. In the results presented in the example analyses of this

thesis, only one of the general mediation methods (structural transformation) provided estimates that were statistically significantly different from both of the traditional mediation methods. However, both of the estimates resulted from general mediation methods (structural transformation and IPW MSMs), were statistically similar. This "triangulation" of estimates when using different approaches may further support evidence that the way that general mediation methods are accounting for the underlying causal model is substantially impacting the estimates of CDM and proportion explained. These differing estimates could represent more severe violation of the assumptions required in order to estimate causal effects when using traditional mediation methods. In the case presented in this thesis, these assumptions could have easily been violated, due to the hypothetical effects that race/ethnicity have on a multitude of potential confounders between payment method (mediator) and low birth weight (outcome). Had these estimates been similar to those resulting from the application of the traditional methods, we may have had more evidence to assume that the required assumptions were not in violation (i.e., there was no relationship with the exposure and confounder(s) of the mediator-outcome association). Thus, using various mediation approaches to address the same question could elucidate on the underlying causal mechanisms involved and allow for the most appropriate estimates to be presented as causal effects.

4.1 LIMITATIONS AND STRENGTHS

As complex as the issues involved in health disparities are, so are the complexities of conducting a review of health disparities literature. For example, the targeted review in this thesis was subject to the limitation that many studies did not explicitly identify themselves as

mediation analyses. These studies met inclusion criteria based on their analysis goals to determine the impact of modifiable intermediate variables on a racial/ethnic disparity in a health outcome, and in effect they were using analytic approaches to answer questions of mediation, but we cannot know if authors were purposefully intending to determine mediation effects. Additionally, while literature on potential issues and required assumptions concerning traditional mediation methods had been published before the timeframe of studies meeting inclusion criteria,[14,50] the publication of many impactful works concerning newly developed issues with traditional mediation approaches in social sciences literature occurred during this timeframe began.[9,51-55] Given the time required for knowledge dissemination, many of the authors of studies included in this review would not have been expected to know the ramifications of using traditional mediation methods to explain disparities based on complex socio-demographic variables.

To our knowledge, no methodologic reviews targeting traditional and general mediation techniques in questions of health disparities have been conducted. While many of the studies included in this review did not explicitly state that they were conducting a mediation analysis, studies were in effect applying analytic techniques to answer questions of mediation. Therefore, these studies are important to include in a methodologic review, as the conclusions drawn from these studies would be similar to conclusions drawn from studies with a stated goal of mediation analysis (i.e., the racial/ethnic disparity would change by a certain magnitude if the proposed mediator(s) could be set at a certain level).

The example analyses were limited by data available on singleton livebirths in the NSFG. Interviews with mothers were subject to recall bias, as 68.5% of births had occurred more than five years previous to the interview and 38.7% of interviews had occurred more than 10 years

previous to the interview. In addition, the sample was limited to births to non-Hispanic White and non-Hispanic Black mothers race/ethnicities for purposes of interpretation, but comparisons are similarly limited to NHW and NHB mothers. While these limitations of the mediation analyses are important to consider, the primary goal of Aim 2 was to demonstrate the range of estimates resulting from different mediation approaches applied to answer the same research question using the same data.

There are multiple strengths to using NSFG data to examine racial/ethnic disparities in birth outcomes. Applying survey weights to NSFG survey data allows for a sample that is representative of births to women living in the United States. The NSFG oversamples non-Hispanic Black women, allowing for more precise estimates of the exposure, outcome, and covariates included in the Aim 2 analysis. The NSFG particularly focuses on factors that are pertinent in questions of birth outcomes, such as gestational length, prenatal care, and delivery payment method.

4.2 CONCLUSION

Social epidemiologists are well-aware of the importance of studying health disparities patterned along non-modifiable socio-demographic factors such as race/ethnicity. It is crucial to examine the impact of modifiable mediators in such disparities in order to understand the nature of such disparities and potentially develop interventions with the end goal of reducing these disparities. However, traditionally-used mediation approaches rely on underlying assumptions that are almost certainly not held when considering complex socio-demographic factors such as race/ethnicity. Recently, advanced analytic techniques have been applied to questions of

mediation in order to develop approaches that can overcome some of the required assumptions of traditional methods. Consideration of underlying assumptions and the use of appropriate mediation approaches can lead to a better understanding of health disparities and potentially more effective interventions.

APPENDIX A: TABLES

Table 1. Characteristics of Simple Random Sample of 50 Articles

Mediation Approach Used	
Difference Method	49 (98%)
Product Method	1 (2%)
General Methods*	0 (0%)
Present Estimate of Proportion of Disparity Explained	7 (14%)
Used Language Inferring Causal Interpretation	43 (86%)
Explicitly Addressed Assumptions Required to Assume Causal Interpretation	0 (0%)

*G-estimation of structural nested models, targeted minimum loss-based estimation, and inverse probability weighted marginal structural models

Table 2. Unweighted Characteristics of National Survey of Family Growth Singleton Livebirths and Low Birth Weight (LBW) Risk

Variable	Total (n = 103,919,318)	%	LBW Events (n = 8,048,388)	LBW Risk	LBW Risk Difference	LBW Risk Ratio
Race/Ethnicity						
Non-Hispanic White	80,090,931	77.1%	5,234,867	0.07	(ref)	(ref)
Non-Hispanic Black	23,828,387	22.9%	2,813,521	0.12	0.05	1.81
Method of Payment						
Medicaid/Govt Assistance	38,092,035	36.7%	3,964,187	0.10	0.04	1.68
Other Forms of Payment	65,827,284	63.3%	4,084,201	0.06	(ref)	(ref)
Maternal Age at Birth						
Under 20 years	24,483,003	23.6%	1,273,812	0.05	(ref)	(ref)
20-24 years	28,817,465	27.7%	1,767,048	0.06	0.01	1.18
25-29 years	32,846,750	31.6%	2,867,579	0.09	0.04	1.68
30-44 years	17,772,101	17.1%	2,139,949	0.12	0.07	2.31
Maternal Education						
< 12th Grade	14,504,697	14.0%	1,622,856	0.11	0.04	1.56
12th Grade or Higher	89,414,621	86.0%	6,425,532	0.07	(ref)	(ref)
Marital Status at Conception						
Married	54,433,700	52.4%	2,856,215	0.05	(ref)	(ref)
Not Married	49,485,619	47.6%	5,192,173	0.10	0.05	2.00
Prenatal Care						
Yes (before 20th week of gestation)	99,187,703	95.4%	7,532,018	0.08	(ref)	(ref)
No	4,731,615	4.6%	516,370	0.11	0.03	1.44
Gestational Length						
<36 weeks	8,525,261	8.2%	4,833,582	0.57	0.53	16.82
36+ weeks	95,394,057	91.8%	3,214,806	0.03	(ref)	(ref)
Wantedness of Pregnancy						
Right Time/Overdue/ Mistimed/Don't Know	88,442,427	85.1%	6,432,941	0.07	(ref)	(ref)
Unwanted	15,476,892	14.9%	1,615,447	0.10	0.03	1.44
Born in United States						
Yes	99,095,551	95.4%	7,502,183	0.08	(ref)	(ref)
No	4,823,768	4.6%	546,205	0.11	0.04	1.50

Table 3. Weighted Counterfactual Disparity Measure (CDM) and Proportion Explained by Payment Method for the Low Birth Weight (LBW) Risk Difference

	Risk Difference	
Unadjusted Disparity	5.27 (3.65, 6.89)	
Mediation Methods	**CDM[a] (95% Confidence Interval)**	**Proportion Explained**
Difference	2.20 (0.64, 3.70)	58.8%
Generalized Product	1.62 (-0.44, 3.68)	69.2%
IPW MSMs*	3.61 (1.46, 5.75)	31.6%
Structural Transformation	4.66 (3.05, 6.28)	11.5%

*IPW MSMs: Inverse Probability-Weighted Marginal Structural Methods

[a]Adjusted for maternal age, maternal education, marital status, prenatal care, gestational length, wantedness of pregnancy, and born in the United States

APPENDIX B: FIFTY STUDIES INCLUDED IN TARGETED REVIEW

First Author	Year	Article Title
Beck[37]	2014	Racial disparities in self-rated health: trends, explanatory factors, and the changing role of socio-demographics
Betancourt[56]	2013	Racial differences in glucose control among patients with type 2 diabetes: a survey on dietary temptations, coping, and trust in physicians
Bhatnagar[57]	2012	Genetic variants in platelet factor 4 modulate inflammatory and platelet activation biomarkers
Cassidy-Bushrow[58]	2012	Race-specific relationship of birth weight and renal function among healthy young children
Celeste[59]	2013	The role of potential mediators in racial inequalities in tooth loss: the Pro-Saude study
Choi[60]	2003	Mechanisms of racial inequalities in prevalence of diarrhoea in South Africa
Cromwell[61]	2005	Race/ethnic disparities in utilization of lifesaving technologies by Medicare ischemic heart disease beneficiaries
De Franco[62]	2016	Racial disparity in previable birth
De Hoog[63]	2011	Overweight at age 2 in a multi-ethnic cohort (ABCD study): the role of prenatal factors, birth outcomes and postnatal factors

Deutscher[64]	2010	Black-white disparities in motor function outcomes taking into account patient characteristics, nontherapy ancillaries, therapy activities, and therapy interventions
Do[38]	2012	Does SES explain more of the black/white health gap than we thought? Revisiting our approach to understanding racial disparities in health
Erickson[65]	2011	The effect of race and ethnicity on outcomes among patients in the ICU: A comprehensive study involving socioeconomic status and resuscitation preferences
Finch[66]	2000	Racial/ethnic disparities in infant mortality: the role of behavioural factors
Flynn[67]	2015	Attribution and emotions regarding health care mistreatment impact on continuity of care among Latino and Anglo America women
Font[68]	2012	Examining the racial disproportionality in child protective services case decisions
Freeman[69]	2011	Association of census tract-level socioeconomic status with disparities in prostate cancer-specific survival
Hamrick[70]	2006	Health care disparities in postmenopausal women referred for DXA screening
Howell[71]	2008	Peginterferon pharmacokinetics in African American and Caucasian American patients with hepatitis C virus genotype 1 infection
Jolly[72]	2010	Cardiac procedures among American Indians and Alaska Natives compared to non-Hispanic Whites hospitalized with ischemic heart disease in California
Karriker-Jaffe[73]	2016	Understanding associations between neighborhood socioeconomic status and negative consequences of drinking: a moderated mediation analysis
Katz[74]	2016	Poverty, Depression, or Lost in Translation? Ethnic and Language Variation in Patient-Reported Rheumatoid Arthritis
Keating[75]	2016	Location isn't everything: proximity, hospital characteristics, choice of hospital, and disparities for breast cancer surgery patients

Kershaw[76]	2011	Metropolitan-level racial residential segregation and black-white disparities in hypertension
King[77]	2015	Decomposing differences in medical care access among cancer survivors by race and ethnicity
Levine[78]	2011	Racial/Ethnic disparities in access to physician care and medications among US stroke survivors
Levine[79]	2014	Acute infection contributes to racial disparities in stroke mortality
Lorch[80]	2007	Racial differences in the use of respiratory medications in premature infants after discharge from the neonatal intensive care unit
Lotoala[81]	2014	Health and wellbeing of older Pacific Peoples in New Zealand
Lynch[82]	2016	Race, ethnicity, psychosocial factors, and telomere length in a multicenter setting
Menon[83]	2007	Amniotic fluid interleukin-1B and interleukin-8 concentrations: racial disparity in preterm birth
Misialek[84]	2014	Socioeconomic status and the incidence of atrial fibrillation (AF) in whites and blacks: the atherosclerosis risk in communities (ARIC) study
Molina[85]	2008	African American and Poor Patients Have a Dramatically Worse Prognosis for Head and Neck Cancer
Moran[86]	2012	Breast Cancers in US Residing Indian-Pakistani vs non-Hispanic White women: comparative analysis of clinical-pathological features, treatment, and survival.
Mukamel[87]	2006	Referrals to high-quality cardiac surgeons: patients' race and characteristics of their physicians
Myakovsky[88]	2012	Perceived discrimination predicts longer time to be accepted for kidney transplant

Osborne[89]	2009	Explaining racial disparities in mortality after abdominal aortic aneurysm repair
Patel[90]	2015	Racial differences in receipt of chlamydia testing among Medicaid-insured women in 2013
Randall[91]	2013	Disparities in revascularization rates after acute myocardial infarction between aboriginal and non-Aboriginal people in Australia
Rangel[92]	2005	Racial and ethnic disparities in influenza vaccination among elderly adults
Runarsdottir[39]	2015	Ethnic differences in youth well-being: The role of sociodemographic background and social support
Ryu[93]	2013	What factors explain disparities in mammography rates among Asian-American immigrant women? A population based study in California
Sambamoorthi[94]	2003	Racial, ethnic, socioeconomic, and access disparities in the use of preventative services among women
Sluyter[95]	2011	Body mass index and percent body fat in a New Zealand multi-ethnic adolescent population
Strumpf[96]	2011	Racial/ethnic disparities in primary care: the role of physician-patient concordance
Taioli[97]	2016	Racial disparities in esophogeal cancer survival after surgery
Tewari[98]	2009	Effect of socioeconomic factors on long-term mortality in men with clinically localized prostate cancer
Valdez[99]	2015	Racial ethnic and socioeconomic disparities in mental health in Arizona
Waddell[100]	2010	Pregnancy risk among black, white, and Hispanic teen girls in New York City public schools

Wang[101]	2009	Reduced rates of primary joint replacement for osteoarthritis in Italian and Greek migrants to Australia: the Melbourne Collaborative Cohort Study
Warner[102]	2010	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

BIBLIOGRAPHY

1. LaVeist TA. *Minority populations and health : an introduction to health disparities in the United States.* 1st ed. San Francisco: Jossey-Bass; 2005.

2. McMichael AJ. Prisoners of the proximate: loosening the constraints on epidemiology in an age of change. *Am. J. Epidemiol.* May 15 1999;149(10):887-897.

3. Harper S, Strumpf EC. Social epidemiology: questionable answers and answerable questions. *Epidemiology.* Nov 2012;23(6):795-798.

4. US Department of Health and Human Services. The Secretary Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020. Phase I report: Recommendations for the framework and format of Healthy People 2020 [Internet}. Section IV: Advisory Committee findings and recommendations [cited 2010 January 6]. Available from: http://www.healthypeople.gov/sites/default/files/PhaseI_0.pdf.

5. Crimmins EM, Kim JK, Alley DE, Karlamangla A, Seeman T. Hispanic paradox in biological risk profiles. *Am. J. Public Health.* Jul 2007;97(7):1305-1310.

6. Peek ME, Cargill A, Huang ES. Diabetes Health Disparities: A Systematic Review of Health Care Interventions. *Medical care research and review : MCRR.* 2007;64(5 Suppl):101S-156S.

7. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers. Soc. Psychol.* Dec 1986;51(6):1173-1182.

8. Pathak EB. Mortality Among Black Men in the USA. *Journal of racial and ethnic health disparities.* Feb 24 2017.

9. VanderWeele T, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface.* 2009;2:457-468.

10. Elwert F, Winship C. Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. *Annual Review of Sociology.* 2014/07/30 2014;40(1):31-53.

11. Naimi AI, Kaufman JS, Howe CJ, Robinson WR. Mediation considerations: serum potassium and the racial disparity in diabetes risk. *The American Journal of Clinical Nutrition.* August 1, 2011 2011;94(2):614-616.

12. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol. Methods.* 02/04 2013;18(2):137-150.

13. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology.* Jan 2009;20(1):18-26.

14. Robins JM. Testing and estimation of direct effects by reparameterizing directed acyclic graphs with structural nested models. *Computation, causation, and discovery.* 1999:349-405.

15. van der Laan MJ, Gruber S. Targeted minimum loss based estimation of causal effects of multiple time point interventions. *The international journal of biostatistics.* 2012;8(1).

16. Robins JM. The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. *Health service research methodology: a focus on AIDS.* 1989;113:159.

17. Robins JM, Mark SD, Newey WK. Estimating exposure effects by modelling the expectation of exposure conditional on confounders. *Biometrics.* 1992:479-495.

18. Van der Laan MJ, Rose S. *Targeted learning: causal inference for observational and experimental data.* Springer Science & Business Media; 2011.

19. Naimi AI, Schnitzer ME, Moodie EE, Bodnar LM. Mediation Analysis for Health Disparities Research. *Am. J. Epidemiol.* Aug 15 2016;184(4):315-324.

20. Warner ET, Tamimi RM, Hughes ME, et al. Racial and Ethnic Differences in Breast Cancer Survival: Mediating Effect of Tumor Characteristics and Sociodemographic and Treatment Factors. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* Jul 10 2015;33(20):2254-2261.

21. Ojha RP, Stallings-Smith S, Flynn PM, Adderson EE, Offutt-Powell TN, Gaur AH. The Impact of Vaccine Concerns on Racial/Ethnic Disparities in Influenza Vaccine Uptake Among Health Care Workers. *American journal of public health.* Sep 2015;105(9):e35-41.

22. Molina Y, Silva A, Rauscher GH. Racial/Ethnic Disparities in Time to a Breast Cancer Diagnosis: The Mediating Effects of Health Care Facility Factors. *Med. Care.* Oct 2015;53(10):872-878.

23. Chen HY, Chauhan SP, Rankins NC, Ananth CV, Siddiqui DS, Vintzileos AM. Racial and ethnic disparities in infant mortality in the United States: the role of gestational age. *Am. J. Perinatol.* Jun 2013;30(6):469-475.

24. Mochari-Greenberger H, Liao M, Mosca L. Racial and ethnic differences in statin prescription and clinical outcomes among hospitalized patients with coronary heart disease. *Am. J. Cardiol.* Feb 1 2014;113(3):413-417.

25. DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. *CA. Cancer J. Clin.* Jan-Feb 2014;64(1):52-62.

26. Jackson CS, Oman M, Patel AM, Vega KJ. Health disparities in colorectal cancer among racial and ethnic minorities in the United States. *J Gastrointest Oncol.* Apr 2016;7(Suppl 1):S32-43.

27. Operario D, Lee JH, Kuo C, Zaller N. Racial and Ethnic Disparities in HIV and STIs in the United States--National Health and Nutrition Examination Survey 1999-2012. *AIDS Patient Care STDS.* Dec 2015;29(12):635-638.

28. Kochanek KD, Anderson RN, Arias E. Leading Causes of Death Contributing to Decrease in Life Expectancy Gap Between Black and White Populations: United States, 1999-2013. *NCHS Data Brief.* Nov 2015(218):1-8.

29. Zhao Y, Kershaw T, Ettinger AS, Higgins C, Lu MC, Chao SM. Association Between Life Event Stressors and Low Birth Weight in African American and White Populations: Findings from the 2007 and 2010 Los Angeles Mommy and Baby (LAMB) Surveys. *Matern. Child Health J.* Oct 2015;19(10):2195-2205.

30. Collins JW, Jr., David RJ. The differential effect of traditional risk factors on infant birthweight among blacks and whites in Chicago. *Am. J. Public Health.* Jun 1990;80(6):679-681.

31. Fuller KE. Low birth-weight infants: the continuing ethnic disparity and the interaction of biology and environment. *Ethn. Dis.* Autumn 2000;10(3):432-445.

32. Hamilton BE, Martin JA, Osterman MJ, Curtin SC, Matthews TJ. Births: Final Data for 2014. *Natl. Vital Stat. Rep.* Dec 2015;64(12):1-64.

33. Dennis JA, Mollborn S. Young maternal age and low birth weight risk: An exploration of racial/ethnic disparities in the birth outcomes of mothers in the United States. *The Social science journal.* Dec 2013;50(4):625-634.

34. Strutz KL, Richardson LJ, Hussey JM. Selected preconception health indicators and birth weight disparities in a national study. *Womens Health Issues.* Jan-Feb 2014;24(1):e89-97.

35. Loggins Clay S, Andrade FC. Role of stress in low birthweight disparities between black and white women: a population-based study. *J. Paediatr. Child Health.* Apr 2015;51(4):443-449.

36. Myaskovsky L, Doebler DA, Posluszny DM, et al. Perceived Discrimination Predicts Longer Time to be Accepted for Kidney Transplant. *Transplantation.* 2012;93(4):423-429.

37. Beck AN, Finch BK, Lin S-F, Hummer RA, Masters RK. Racial disparities in self-rated health: trends, explanatory factors, and the changing role of socio-demographics. *Social Science & Medicine.* 2014;104:163-177.

38. Do DP, Frank R, Finch BK. Does SES explain more of the black/white health gap than we thought? Revisiting our approach toward understanding racial disparities in health. *Soc Sci Med.* May 2012;74(9):1385-1393.

39. Runarsdottir EM, Vilhjalmsson R. Ethnic differences in youth well-being: The role of sociodemographic background and social support. *Scand. J. Public Health.* 2015:1403494815588644.

40. de Hoog ML, van Eijsden M, Stronks K, Gemke RJ, Vrijkotte TG. Overweight at age two years in a multi-ethnic cohort (ABCD study): the role of prenatal factors, birth outcomes and postnatal factors. *BMC Public Health.* Aug 01 2011;11:611.

41. Osborne NH, Upchurch GR, Jr., Mathur AK, Dimick JB. Explaining racial disparities in mortality after abdominal aortic aneurysm repair. *J. Vasc. Surg.* Oct 2009;50(4):709-713.

42. Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. Alternative Definitions of "Proportion Eliminated". *Epidemiology (Cambridge, Mass.).* 2014;25(2):308-309.

43. Wall MM, Poh E, Cerdá M, Keyes KM, Galea S, Hasin DS. Commentary on Harper S, Strumpf EC, Kaufman JS. Do Medical Marijuana Laws Increase Marijuana Use? Replication Study and Extension. *Annals of epidemiology.* 04/24 2012;22(7):536-537.

44. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology.* Lippincott Williams & Wilkins; 2008.

45. Rubin DB. Practical implications of modes of statistical inference for causal effects and the critical role of the assignment mechanism. *Biometrics.* Dec 1991;47(4):1213-1234.

46. Pearl J. Interpretation and identification of causal mediation. *Psychol. Methods.* Dec 2014;19(4):459-481.

47. Vansteelandt S. Estimation of Direct and Indirect Effects. *Causality*: John Wiley & Sons, Ltd; 2012:126-150.

48. Loeys T, Moerkerke B, Raes A, Rosseel Y, Vansteelandt S. Estimation of Controlled Direct Effects in the Presence of Exposure-Induced Confounding and Latent Variables. *Structural Equation Modeling: A Multidisciplinary Journal.* 2014/07/03 2014;21(3):396-407.

49. Beyea J, Greenland S. The Importance of Specifying the Underlying Biological Model in Estimating The Probability of Causation. *Health Phys.* 1999;76(3):269-274.

50. Kaufman JS, Cooper RS. Seeking causal explanations in social epidemiology. *Am. J. Epidemiol.* Jul 15 1999;150(2):113-120.

51. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu. Rev. Psychol.* 2007;58:593-614.

52. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol. Methods.* Dec 2010;15(4):309-334.

53. Zhao X, Lynch JG, Chen Q. Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research.* 2010-08-01 00:00:00 2010;37(2):197-206.

54. Rucker DD, Preacher KJ, Tormala ZL, Petty RE. Mediation Analysis in Social Psychology: Current Practices and New Recommendations. *Social and Personality Psychology Compass.* 2011;5(6):359-371.

55. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychol. Methods.* 2013;18(2):137.

56. Betancourt RM, Degnan KO, Long JA. Racial differences in glucose control among patients with type 2 diabetes: a survey on dietary temptations, coping, and trust in physicians. *Ethn. Dis.* Autumn 2013;23(4):409-414.

57. Bhatnagar P, Lu X, Evans MK, et al. Genetic variants in platelet factor 4 modulate inflammatory and platelet activation biomarkers. *Circulation. Cardiovascular genetics.* Aug 01 2012;5(4):412-421.

58. Cassidy-Bushrow AE, Wegienka G, Barone CJ, 2nd, et al. Race-specific relationship of birth weight and renal function among healthy young children. *Pediatr. Nephrol.* Aug 2012;27(8):1317-1323.

59. Celeste RK, Goncalves LG, Faerstein E, Bastos JL. The role of potential mediators in racial inequalities in tooth loss: the Pro-Saude study. *Community Dent. Oral Epidemiol.* Dec 2013;41(6):509-516.

60. Choi SY. Mechanisms of racial inequalities in prevalence of diarrhoea in South Africa. *J. Health Popul. Nutr.* Sep 2003;21(3):264-272.

61. Cromwell J, McCall NT, Burton J, Urato C. Race/ethnic disparities in utilization of lifesaving technologies by Medicare ischemic heart disease beneficiaries. *Med. Care.* Apr 2005;43(4):330-337.

62. DeFranco EA, Hall ES, Muglia LJ. Racial disparity in previable birth. *Am. J. Obstet. Gynecol.* 2016;214(3):394. e391-394. e397.

63. de Hoog ML, van Eijsden M, Stronks K, Gemke RJ, Vrijkotte TG. Overweight at age two years in a multi-ethnic cohort (ABCD study): the role of prenatal factors, birth outcomes and postnatal factors. *BMC Public Health.* 2011;11(1):611.

64. Deutscher D, Horn SD, Smout RJ, DeJong G, Putman K. Black-white disparities in motor function outcomes taking into account patient characteristics, nontherapy ancillaries, therapy activities, and therapy interventions. *Arch. Phys. Med. Rehabil.* Nov 2010;91(11):1722-1730.

65. Erickson SE, Vasilevskis EE, Kuzniewicz MW, et al. The effect of race and ethnicity on outcomes among patients in the intensive care unit: a comprehensive study involving socioeconomic status and resuscitation preferences. *Crit. Care Med.* Mar 2011;39(3):429-435.

66. Finch BK, Frank R, Hummer RA. Racial/ethnic disparities in infant mortality: The role of behavioral factors. *Soc. Biol.* 2000;47(3-4):244-263.

67. Flynn PM, Betancourt H, Garberoglio C, Regts GJ, Kinworthy KM, Northington DJ. Attributions and emotions regarding health care mistreatment impact continuity of care among Latino and Anglo American women. *Cultur. Divers. Ethnic Minor. Psychol.* 2015;21(4):593.

68. Font SA, Berger LM, Slack KS. Examining Racial Disproportionality in Child Protective Services Case Decisions. *Children and youth services review.* Nov 01 2012;34(11):2188-2200.

69. Freeman VL, Ricardo AC, Campbell RT, Barrett RE, Warnecke RB. Association of census tract-level socioeconomic status with disparities in prostate cancer-specific survival. *Cancer Epidemiol. Biomarkers Prev.* Oct 2011;20(10):2150-2159.

70. Hamrick I, Steinweg KK, Cummings D, Whetstone L. Health care disparities in postmenopausal women referred for DXA screening. *FAMILY MEDICINE-KANSAS CITY-.* 2006;38(4):265.

71. Howell CD, Dowling TC, Paul M, et al. Peginterferon pharmacokinetics in African American and Caucasian American patients with hepatitis C virus genotype 1 infection. *Clin. Gastroenterol. Hepatol.* 2008;6(5):575-583.

72. Jolly S, Kao C, Bindman AB, Korenbrot C. Cardiac procedures among American Indians and Alaska Natives compared to non-Hispanic Whites hospitalized with ischemic heart disease in California. *J. Gen. Intern. Med.* 2010;25(5):430-434.

73. Karriker-Jaffe KJ, Liu H, Kaplan LM. Understanding associations between neighborhood socioeconomic status and negative consequences of drinking: a moderated mediation analysis. *Prev. Sci.* 2016;17(4):513-524.

74. Katz PP, Barton J, Trupin L, et al. Poverty, Depression, or Lost in Translation? Ethnic and Language Variation in Patient-Reported Outcomes in Rheumatoid Arthritis. *Arthritis care & research.* 2016;68(5):621-628.

75. Keating NL, Kouri EM, He Y, Freedman RA, Volya R, Zaslavsky AM. Location isn't everything: proximity, hospital characteristics, choice of hospital, and disparities for breast cancer surgery patients. *Health Serv. Res.* 2016.

76. Kershaw KN, Roux AVD, Burgard SA, Lisabeth LD, Mujahid MS, Schulz AJ. Metropolitan-level racial residential segregation and black-white disparities in hypertension. *Am. J. Epidemiol.* 2011:kwr116.

77. King CJ, Chen J, Dagher RK, Holt CL, Thomas SB. Decomposing differences in medical care access among cancer survivors by race and ethnicity. *Am. J. Med. Qual.* 2015;30(5):459-469.

78. Levine DA, Neidecker M, Kiefe CI, Karve S, Williams LS, Allison JJ. Racial/ethnic disparities in access to physician care and medications among US stroke survivors. *Neurology.* 2011;76(1):53-61.

79. Levine DA, Langa KM, Rogers MA. Acute infection contributes to racial disparities in stroke mortality. *Neurology.* 2014;82(11):914-921.

80. Lorch SA, Wade KC, Bakewell-Sachs S, Medoff-Cooper B, Escobar GJ, Silber JH. Racial differences in the use of respiratory medications in premature infants after discharge from the neonatal intensive care unit. *The Journal of pediatrics.* 2007;151(6):604-610. e601.

81. Lotoala F, Breheny M, Alpass F, Henricksen A. Health and wellbeing of older Pacific Peoples in New Zealand. *N. Z. Med. J.* Dec 19 2014;127(1407):27-39.

82. Lynch SM, Peek M, Mitra N, et al. Race, ethnicity, psychosocial factors, and telomere length in a multicenter setting. *PloS one.* 2016;11(1):e0146723.

83. Menon R, Williams SM, Fortunato SJ. Amniotic fluid interleukin-1β and interleukin-8 concentrations: racial disparity in preterm birth. *Reproductive Sciences.* 2007;14(3):253-259.

84. Misialek JR, Rose KM, Everson-Rose SA, et al. Socioeconomic status and the incidence of atrial fibrillation in whites and blacks: the Atherosclerosis Risk in Communities (ARIC) study. *Journal of the American Heart Association.* 2014;3(4):e001159.

85. Molina MA, Cheung MC, Perez EA, et al. African American and poor patients have a dramatically worse prognosis for head and neck cancer. *Cancer.* 2008;113(10):2797-2806.

86. Moran MS, Gonsalves L, Goss DM, Ma S. Breast cancers in U.S. residing Indian-Pakistani versus non-Hispanic White women: comparative analysis of clinical-pathologic features, treatment, and survival. *Breast Cancer Res. Treat.* 02/08 2011;128(2):543-551.

87. Mukamel DB, Weimer DL, Mushlin AI. Referrals to High-Quality Cardiac Surgeons: Patients' Race and Characteristics of Their Physicians. *Health Serv. Res.* 2006;41(4p1):1276-1295.

88. Myaskovsky L, Doebler DA, Posluszny DM, et al. Perceived discrimination predicts longer time to be accepted for kidney transplant. *Transplantation.* 2012;93(4):423.

89. Osborne NH, Upchurch GR, Mathur AK, Dimick JB. Explaining racial disparities in mortality after abdominal aortic aneurysm repair. *J. Vasc. Surg.* 2009;50(4):709-713.

90. Patel CG, Chesson HW, Tao G. Racial Differences in Receipt of Chlamydia Testing Among Medicaid-Insured Women in 2013. *Sex. Transm. Dis.* 2016;43(3):147-151.

91. Randall DA, Jorm LR, Lujic S, O'Loughlin AJ, Eades SJ, Leyland AH. Disparities in revascularization rates after acute myocardial infarction between aboriginal and non-aboriginal people in Australia. *Circulation.* 2013:CIRCULATIONAHA. 112.000566.

92. Rangel MC, Shoenbach VJ, Weigle KA, Hogan VK, Strauss RP, Bangdiwala SI. Racial and ethnic disparities in influenza vaccination among elderly adults. *J. Gen. Intern. Med.* 2005;20(5):426-431.

93. Ryu SY, Crespi CM, Maxwell AE. What factors explain disparities in mammography rates among Asian-American immigrant women? A population-based study in California. *Womens Health Issues.* 2013;23(6):e403-e410.

94. Sambamoorthi U, McAlpine DD. Racial, ethnic, socioeconomic, and access disparities in the use of preventive services among women. *Prev. Med.* 2003;37(5):475-484.

95. Sluyter JD, Schaaf D, Scragg RK, Plank LD. Body mass index and percent body fat in a New Zealand multi-ethnic adolescent population. *International Journal of Pediatric Obesity.* 2011;6(1):36-44.

96. Strumpf EC. Racial/ethnic disparities in primary care: the role of physician-patient concordance. *Med. Care.* 2011;49(5):496-503.

97. Taioli E, Wolf AS, Camacho-Rivera M, et al. Racial disparities in esophageal cancer survival after surgery. *J. Surg. Oncol.* 2016.

98. Tewari AK, Gold HT, Demers RY, et al. Effect of socioeconomic factors on long-term mortality in men with clinically localized prostate cancer. *Urology.* 2009;73(3):624-630.

99. Valdez LA, Langellier BA. Racial/Ethnic and Socioeconomic Disparities in Mental Health in Arizona. *Frontiers in public health.* 2015;3.

100. Waddell EN, Orr MG, Sackoff J, Santelli JS. Pregnancy risk among black, white, and Hispanic teen girls in New York City public schools. *J. Urban Health.* 2010;87(3):426-439.

101. Wang Y, Simpson JA, Wluka AE, et al. Reduced rates of primary joint replacement for osteoarthritis in Italian and Greek migrants to Australia: the Melbourne Collaborative Cohort Study. *Arthritis research & therapy.* 2009;11(3):R86.

102. 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. *J. Community Health.* 2010;35(4):398-408.