2018

Association Between Genetic Ancestry and Body Mass Index Among a Cohort of Hispanic American Children

Sahel Hazrati
Walden University

Walden University

College of Health Sciences

This is to certify that the doctoral dissertation by

Sahel Hazrati

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Namgyal Kyulo, Committee Chairperson, Public Health Faculty
Dr. Scott McDoniel, Committee Member, Public Health Faculty
Dr. Chinaro Kennedy, University Reviewer, Public Health Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2018

Abstract

Association Between Genetic Ancestry and Body Mass Index Among a Cohort of Hispanic American Children

by

Sahel Hazrati

MPH, George Mason University, 2012

BS, Shahid Beheshti University of Medical Sciences, 1995

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

November 2018

Abstract

Childhood obesity is disproportionately higher among children from Hispanic backgrounds. Ethnicity is a social and cultural construct and does not capture true ancestral heterogeneity. Hispanic Americans have a wide variety of genetic admixture proportions of European (EUR), Native American (AMR), and African (AFR) ancestry. The objective of this cross-sectional study was to assess the contribution of ancestral genetic composition to body mass index (BMI), and to evaluate the relationship of obesity risk factors to BMI among 154 2-year-old Hispanic American children. The theory of Evolutionary Developmental Biology was utilized to investigate the relationship between children's growth process and ancestral background. Their genetic admixture was estimated using the ancestry and kinship toolkit and BMI was calculated and evaluated using the Center for Disease Control and Prevention (CDC) BMI charts. Three simple linear regressions assessed the association between standard EUR, AMR, and AFR to BMI. A backward, stepwise, linear regression was performed to evaluate the influence of sex, birth weight, and juice consumption frequency as well as mother's age, BMI, education, and region of birth on the child's BMI. No associations were found between BMI and genetic admixture proportion, and the regression model revealed that only birth weight was positively associated with BMI; higher maternal education was negatively associated with BMI. Contrary to adulthood obesity studies, EUR**,** AMR, and AFR proportions were not associated with BMI at age 2, which suggests that the influence of genetic composition on BMI may vary by age. This information has the potential to create positive social change by developing preventions that target modifiable risk factors, such as maternal education.

Association Between Genetic Ancestry and Body Mass Index Among a Cohort of Hispanic American Children

by

Sahel Hazrati

MPH, George Mason University, 2012

BS, Shahid Beheshti University of Medical Sciences, 1995

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

November 2018

Dedication

I dedicate my PhD dissertation to Ramtin, my one and only son. He has always been my inspiration to work hard and continue my education as in the United States. He motivates me to be enthusiastic and passionate to make world a better place for women and children. Without his love and support I wouldn't be able to accomplish this journey. I also dedicate this dissertation to my parents, my sisters, my brothers and my niece Sarina.

Acknowledgments

I would like to express that I am very grateful to all who supported and guided me during the process of my dissertation. I especially thank my faculty chair, Dr. Namgyal Kyulo, my committee member Dr. Scott McDoniel and my URR Dr. Chinaro Kennedy.

I also express my appreciation and gratitude to the Inova Translational Medicine Institute for allowing me to use data collected for "The First 1,000 Days of Life and Beyond Study" ("The First 1,000 Days of Life and Beyond", 2018).

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

From 1971 to 2012 in the United States, childhood obesity increased from 5.2% to 16.9% among children and adolescents 2–19 years old (Fryar, Carroll, & Ogden, 2012). Childhood obesity is a major health problem in the United States and it is disproportionately higher among Hispanic American children (Isasi, 2016). Childhood obesity occurs when a child's weight is above a level considered healthy for their age and height. According to the Centers for Disease Control and Prevention (CDC), excessive weight gain in children is due to environmental, behavioral, and genetic factors, just as it is for adults (CDC, 2016a). Obese children will more likely develop type II diabetes and cardiovascular diseases and remain obese into adulthood (Sahoo et al., 2015). Some studies have reported higher consumption of less nutritious food and drinks among Hispanic American children as compared to non-Hispanic Whites, such as French fries and sugar-sweetened beverages (Batis, Hernandez-Barrera, Barquera, Rivera, & Popkin, 2011); however, non-overweight Hispanic and overweight Hispanic children have a similar quality diet (National Council of La Raza &Wilson, 2009). Hispanic Americans are the fastest-growing ethnicity in the United States; by 2060, 33.5% of American children will be of Hispanic descent; therefore, addressing the disparity of obesity is important for U.S. healthcare and productivity of country (Colby & Ortman, 2015).

Obesity is a complex disease with multifactorial causation; however, disparities in the prevalence of obesity may demonstrate a substantial genetic component along with an obesogenic environment (Dubois et al., 2012). The genomic composition of Hispanic American is a tri-ethnic genetic admixture of Native American (AMR), European (EUR),

and African (AFR) traits (Hunninghake, Weiss, & Celedón, 2006). The specific proportions of this genetic admixture in Hispanic American children may contribute to the etiology of obesity beyond ethnic classification. Cardel et al. (2011) studied the association between AFR and EUR in the etiology of racial differences in body composition during childhood among children from different reported races, they suggested that EUR is associated with lower lean mass and AFR is associated with lower fat mass and higher bone mineral content. In this study, I investigated the association between genetic admixture proportions and BMI among Hispanic American children from 24 different parental countries of origin in the Washington, DC, metropolitan area. The discovery of any association or no association between genetic ancestry admixture proportion and obesity among Hispanic American children could help healthcare providers develop more efficient policies to prevent and control obesity and to decrease its economic burden in the U.S. healthcare system.

Background

Obesity is an outcome of a positive energy imbalance between energy intake and expenditure; however, recent studies indicate that genetic background is another important obesity risk factor, and that some children are at a higher risk due to genetic factors (Garver et al., 2013; Sahoo et al., 2015). Childhood obesity influences physical and psychological health. Obese children will more likely stay obese and develop other chronic conditions, such as type II diabetes and cardiovascular diseases (Sahoo et al., 2015).

The National Health and Nutrition Examination Survey (NHANES) reported the prevalence of obesity as 21.9% for Hispanic children and 14.7% for non-Hispanic White children, aged 2–19 years. As the number of Hispanic Americans is rapidly growing in the United States, the total number of Hispanic American obese children will continue to increase if all factors are not investigated and addressed. In the NHANES dataset, participants reported their self-identified race and ethnicity (Ogden, Carroll, Kit, & Flegal, 2014). In general, race and ethnicity are misunderstood terms among scientists (Yudell, Roberts, DeSalle, & Tishkoff, 2016). The use of self-identified race and ethnicity categories is not sufficient for characterizing the genetic background of Hispanics (Mersha & Abebe, 2015). Genetic admixture represents a biological aspect of race and ethnicity and has been used to investigate the association between ancestry and risk of diseases. Also, it may help to identify genes that may contribute to certain conditions (Fernandez, Pearson, Kell, & Brown, 2013). The genetic ancestry of Hispanic individuals varies across geographic locations (Bryc, Durand, Macpherson, Reich, & Mountain, 2015). Generally, Hispanics have a varied range of ancestry admixture proportions of EUR, AMR, and AFR (Hunninghake et al., 2006). According to Skotte, Korneliussen, and Albrechtsen, (2013) "Admixture occurs when isolated populations begin interbreeding and their offspring represent a mixture of alleles from different ancestral populations" (p. 693).

Due to new genotyping and sequencing technologies, knowledge about the genetic susceptibility of obesity has increased (Herrera, Keildson, & Lindgren, 2011). Cardel et al. (2011) suggested that the AFR admixture influences levels of adiposity in

young children. Furthermore, the associations between ancestry admixture and other chronic diseases—such as cancer and type 2 diabetes—have been reported in the literature (Divers et al., 2013; Ricks-Santi et al., 2012). The specific admixture proportions of children may contribute to the etiology of childhood obesity among Hispanics. Knowledge about the association between genetic admixture and obesity may help to develop targeted obesity prevention and interventions for Hispanic children from different ancestry backgrounds. Targeted and precision preventions may decrease the incidence of childhood obesity and effectively eliminate disparities of obesity between different groups of American children and the consequences of obesity over time.

Problem Statement

Childhood obesity is a major public health concern in the United States and it is higher among ethnic minorities. According to Ogden et al. (2014), 29.8% and 16.7% of Hispanic American preschool children were overweight and obese, respectively; this is twice the national average for this age group. Obesity is a complex disease, and many studies have contributed to our understanding of the major risk factors of obesity, such as dietary intake and physical activity. Excessive exposure to unhealthy food and limited or no access to healthy options, such as healthy grocery stores in disadvantaged neighborhoods, is related to the high prevalence of obesity seen in minorities living in these communities (Ellaway, Anderson, & Macintyre, 1997; Kahn, Tatham, Pamuk, & Heath, 1998). Furthermore, epidemiologic studies argue that intrauterine exposure to smoking and higher than recommended weight gain during pregnancy may increase glucose levels and result in early onset obesity; too much screen time including TV and

video games and short sleep duration may also increase risks for obesity (Taveras, Rifas-Shiman, Oken, Gunderson, & Gillman, 2008).

The prevalence of childhood obesity among Hispanic American children is partially explained by socioeconomic and cultural factors (Taveras, Gillman, Kleinman, Rich-Edwards, Rifas-Shiman, 2013); however, knowledge about disparities within ethnicities is not well investigated. Hispanic American children's parents are originally from South or Central American countries, with a wide variety of admixture proportions of EUR, AMR, and AFR ancestry. An association of West African, EUR and AMR ancestry and adult body mass index (BMI) has been reported in the United States (Fernandez et al., 2013; Shaffer et al., 2007). Cardel et al. (2011) suggested that genetic factors may contribute to total body fat accumulation among children. Nevertheless, the association between genetic admixture proportion and obesity among Hispanic American children has not been investigated. Ancestral genetic background may explain some differences in the prevalence of obesity among Hispanic American children, with similar obesogenic factors and socioeconomic status (Fernandez et al., 2013; Higgins, Fernández, Goran, & Gower, 2005).

Purpose of the Study

The purpose of this quantitative, cross-sectional research study was to investigate the contribution of children's genetic admixture proportion—including EUR, Native AMR and AFR—to childhood obesity through the use of BMI among 2-year-old Hispanic American children. I also investigated whether this association remained significant after controlling for obesogenic risk factors of early life, including frequency

of daily juice consumption at age 2, sex, birth weight. maternal education, and household income.

The independent variable in this study was the admixture proportion (EUR, AMR, and AFR). The dependent variable was BMI at age 2. Potential confounding variables were as follows: (a) frequency of juice consumption at age 2; (b) sex; (c) maternal education, (d) mother's BMI (e) mother's age (f) mother's region of birth and (g) birth weight (Brophy, 2009; Sahoo et al., 2015).

Research Questions and Hypotheses

This study was guided by three research questions.

RQ1: Is there an association between children's EUR genetic background and BMI among 2-year-old Hispanic American children?

H_{01}: There is no statistically significant association between children's EUR genetic background and BMI among Hispanic American children.

H_{A1}: There is a statistically significant association between children's EUR genetic background and BMI among Hispanic American children.

To examine RQ1, simple linear regression analysis was conducted to assess whether EUR genetic background of children influences BMI. If an association was found, then multiple linear regression would have been conducted to determine whether this association remained significant even after controlling for confounding variables.

RQ2: Is there an association between AMR genetic background and BMI among 2-year-old Hispanic American children?

H_{02}: There is no statistically significant association between children's AMR genetic background and BMI among Hispanic American children.

H_{A2}: There is a statistically significant association between children's AMR genetic background and BMI among Hispanic American children.

To examine RQ2, simple linear regression analysis was conducted to assess whether AMR genetic background of children is associated with BMI. If an association is found, then multiple linear regression would have been conducted to determine whether this association remained even after controlling for confounding variables.

RQ3: Is there an association between children's AFR genetic background and BMI among 2-year-old Hispanic American children?

H_{03}: There is no statistically significant association between children's AFR genetic background and BMI among Hispanic American children.

H_{A3}: There is a statistically significant association between children's AFR genetic background and BMI among Hispanic American children.

To examine RQ3, simple linear regression analysis was conducted to determine whether AFR genetic background of children was associated with BMI. If an association was found, then multiple linear regression would have been conducted to determine whether this association remained significant even after controlling for confounding variables.

Theoretical Framework for the Study

The theory of evolutionary developmental biology (evo-devo) was used to investigate the relationship between children's growth process and ancestral background.

Ancestral genetic background may explain some differences in the prevalence of obesity among Hispanic American children with similar social and environmental risk factors. Precise understanding of association of ancestry background of Hispanic children and the prevalence of obesity may help to develop more precise and personalized prevention and intervention policies.

The genetic background of AFR, AMR, and EUR ancestry is associated with the body composition of adults in the United States (Klimentidis, Miller, & Shriver, 2009). Similar to many other individuals in the United States, Hispanic Americans have this kind of admixed ancestry. According to Salzano and Sans (2014), "Latin American populations can be viewed as natural experiments for the investigation of unique anthropological and epidemiological issues." Evo-devo explains the relationship between growth process and ancestral background. The Life-History Theory (Hill, 1993) is an evo-devo domain that provides an important conceptual framework to address questions about health and disease (Gluckman, Low, Buklijas, Hanson, & Beedle, 2011).

Many features of human anatomy are related to consequences of evolutionary history. The principles of evolutionary biology may provide new insights into childhood obesity and enable an integrated understanding of human biology (Gluckman et al., 2011). Epigenetic findings have increased knowledge of the molecular mechanisms that may contribute to the programming of obesity (Kappil, Wright, & Sanders, 2016). These concepts are explored in more detailed analysis in Chapter 2.

Nature of the Study

This is a cross-sectional study investigating the contribution of children's genetic admixture proportion to childhood obesity through use of BMI among 2-year-old Hispanic American children, adjusted for major obesogenic risk factors of early life obesity such as juice consumption and socioeconomic risk factors. The data for my study was obtained from the Inova Translational Medicine Institute. Inova's "The First 1,000 Days of Life and Beyond" is a genomic longitudinal cohort study based in the Inova Health System in the Washington, DC metropolitan area. R 3.1.2 and SAS 9.4 were used to conduct analyses of descriptive and inferential statistics. This is further addressed in Chapter 3.

Operational Definitions

Dependent Variables

BMI: BMI is the ratio of person's body weight in kilograms to person's height square. BMI is used to classify and assess individuals as overweight and obese. BMI can estimate how much body fat a person has (Nuttall, 2015).

Obesity: For children and adolescents 2–19 years old, obesity is defined as having a BMI at or above the 95th percentile of the sex-specific BMI on the CDC Age growth chart (CDC, 2016b).

Overweight: For 2–19 years old children and adolescents "overweight" is defined as having a BMI at or above the 85th percentile of the sex-specific BMI on the CDC Age growth chart (CDC, 2016b).

Independent Variable

Genetic ancestry admixture proportion: "Admixture occurs when isolated populations begin interbreeding and their offspring represent a mixture of alleles from different ancestral populations. Estimating the admixture proportions of an individual is a valuable tool in both population genetics and genetic epidemiology" (Skotte et al., 2013, p. 693). The ancestry of study children was estimated by the ancestry and kinship toolkit (AKT; Arthure, Schulz-Trieglaff, Cox, & O'Connell, 2016) using 17,535 reliable and common single-nucleotide polymorphisms (SNPs) by projecting the samples into the 1,000 genomes (Auton et al., 2015) principal components (PCs), followed by assigning the PCs to admixture proportions.

Admixture Proportions Super Groups

- *EUR:* European,
- *AFR:* African,
- *AMR:* Americas,
- *SAS:* South Asia, and
- *EAS:* East Asia.

Confounding Variables

- Juice consumption frequency at age 2,
- Maternal education,
- Maternal age,
- Maternal BMI,
- Maternal region of birth,

- Birth weight, and
- Sex

Other Terms

Obesogenic factors: factors that contribute to obesity such as high-calorie food intake and physical inactivity.

Epigenetic factors: Epigenetics is the study of chemical reactions that activate or deactivate parts of the genome in certain times and locations (Genetic Science Learning Center, 2018).

Evo-devo: According to Hall (2012), "Evolutionary developmental biology (evo-devo) as a discipline is concerned, among other things, with discovering and understanding the role of changes in developmental mechanisms in the evolutionary origin of aspects of the phenotype" (p. 184). In a very real sense, evo-devo opens the black box between genotype and phenotype, or more properly, phenotypes as multiple life history stages arise in many organisms from a single genotype (Hall, 2012).

Genome-wide association study (GWAS): According to the National Human Genome Research Institute:

> A genome-wide association study is an approach that involves rapidly scanning markers across complete sets of DNA, or genomes, of many people to find genetic variations associated with a particular disease. Once new genetic associations are identified, researchers can use the information to develop better strategies to detect, treat and prevent the disease. Such studies are particularly useful in finding genetic variations that contribute to common, complex diseases,

such as asthma, cancer, diabetes, heart disease and mental illnesses. (National Human Genome Research Institute, 2015)

Assumptions

The study population included about 150 2-year-old Hispanic American children who were recruited into "The First 1,000 Days of Life and Beyond" study (The First 1,000 Days of Life and Beyond, 2018). I assumed that the sample was representative of Hispanic American children in the Washington, DC metropolitan area. For this research, I used secondary data, and I assumed that the parentally reported anthropometrics were accurate, and that other reported data were truthful. I also assumed that the secondary data were retrieved correctly. I assumed that the AKT used in bioinformatics department is a reliable, statistical genetics tool for analyzing large cohorts of whole-genome sequenced samples.

Scope and Delimitations

This quantitative study evaluated 2-year-old Hispanic American children who were recruited into "The First 1,000 Days of Life and Beyond," a genomic study at Inova Health System from July 2012 to 2015 in Falls Church, Virginia. The goal of this study was to investigate the contribution of children's genetic admixture proportion (*EUR, AMR, and AFR*) to childhood obesity (determined via BMI) for 2-year-old Hispanic American children. Furthermore, major obesogenic risk factors of early life obesity such as juice consumption and socioeconomic risk factors were evaluated. Hispanics have a varied range of ancestry admixture proportions of EUR, AMR, and AFR (Hunninghake, 2006). When an isolated population starts interbreeding, their offspring represents a

mixture of alleles from various ancestries; this constitutes admixture. Estimating the admixture proportions of an individual is valuable in both population genetics and genetic epidemiology. In population genetics, admixture analysis allows the researcher to classify individuals with unknown races into discrete populations. This has been used to successfully describe the genetics of different populations and even extinct populations (Rasmussen, 2010; Rosenberg et al., 2002). Knowing individual admixture proportions is also useful in genetic association studies.

The objective of study objective was chosen to investigate how a mixture of alleles from different ancestral populations contributes to childhood obesity regardless of cultural and social risk factors. I considered all the available confounding variables including prenatal, clinical, socioeconomic, and early life dietary variables to improve internal validity. Anthropometric values were used to calculate age-sex-specific BMI; weights and heights were parentally reported; I used the interquartile outlier rule to detect and remove outliers (Hazrati et al., 2016). Furthermore, I removed all the physiologically impossible parentally reported anthropometrics. I believe Hispanic American children in my study were representative of Hispanic American children in the Washington, DC metropolitan area. In general, retrospective studies can be riddled with threats to both internal and external validity. Although a cause-and-effect relationship cannot be determined using retrospective studies, they are useful for providing preliminary data and in guiding the development of future prospective studies (Tofthagen, 2012).

Limitations

This was a cross-sectional study using parentally reported anthropometrics; some of the confounding variables—including dietary values—are parentally reported. Therefore, differential misclassification and information bias may have occurred during data collection, and internal validity could be jeopardized. Another threat to validity of this study is participation bias: firstly, parents who agreed to participate in the genomic study may have had different demographics than the ones who did not agree to participate in a longitudinal genomic study; secondly, parents who did not complete longitudinal surveys may have had a particular problem affecting this cross-sectional study. This is a particular problem when the characteristics of non-responders differ from responders (Shepherd, Power, & Carter, 1998).

This study was limited to 2-year-old Hispanic American children residing in the Washington, DC metropolitan area who were recruited into "The First 1,000 Days of Life" study and whose parents had completed longitudinal surveys. Although the children in this study are likely representative of Hispanic American ancestry admixture in the Washington, DC metropolitan area, the result may not be generalizable to different acculturation and assimilation processes in other regions of the United States.

Significance

Hispanic children have a higher prevalence of obesity compared to non-Hispanic White children in the United States (Ogden et al., 2014). However, knowledge about association of obesity and children's ancestry genetic background is lacking. The use of mother's self-identified Hispanic ethnicity is not sufficient to explain the disparities in

childhood obesity, and this issue has not been well studied among Hispanic American children from different ancestries. Very little is known about how admixture proportions may affect childhood obesity. A study of the associations between obesity and genetic admixture proportions of Hispanic American children—along with other social and clinical factors—could help to better understand the higher prevalence and etiology of childhood obesity among Hispanic children.

Social Change Implications

If the prevalence or severity of childhood obesity among Hispanic children is related to specific ancestry background, then knowledge of this association is beneficial; it can help healthcare providers develop more effective and practical public health policies to prevent and treat childhood obesity for this population. The information gained from this study has the potential to create positive social change by developing precision preventions. The knowledge may help to develop targeted obesity prevention interventions for Hispanic children from different ancestry backgrounds. Targeted and precision prevention may decrease childhood obesity and more effectively eliminate the disparity of obesity and lower the consequences of obesity over time.

Summary

In this chapter I have provided some background on the prevalence and disparities in childhood obesity in the United States. Childhood obesity is a main public health issue, and it is disproportionately higher for Hispanic American children. Hispanic Americans are the fastest growing minority group in the United States, and it seems high prevalence of childhood obesity will continue to be an issue if all the risk factors of obesity are not

investigated precisely. As stated before, in the collection of national health data such as NHANES, people reported their self-identified and socially assigned race or ethnicity. Hispanic Americans have an admixed ancestry background and use of self-reported ethnicity is not sufficient to investigate the biological aspects? of their background. As theory of evo-devo explains the relationship between growth process and ancestral background, I intend to use genetic ancestry to investigate the association between the biological aspect of children's backgrounds and obesity. I used a genetic admixture of Hispanic American children as a surrogate for the biological aspect of their ancestry. A precise understanding of the association between ethnic background and obesity could help to develop more precise and personalized prevention and treatment policies.

In Chapter 2, I briefly explain the literature search strategy and reviewe studies that gave more background on the issues of childhood obesity and disparities in obesity. I discuss the theoretical framework and the economic impact of childhood obesity. I also argue the insufficiency of the social constructs of race and ethnicity in scientific study.

Chapter 2: Literature Review

Childhood obesity is a major public health concern in the United States, and it is higher among ethnic minorities. According to Ogden et al. (2014), 29.8% and 16.7% of Hispanic American preschool children were overweight and obese, respectively. This is twice the national average for this age group. Racial and ethnic disparities in childhood obesity among Hispanic American children are partially explained by socioeconomic and cultural factors; however, these disparities within single ethnicities are not well investigated. Hispanic American children's parents are originally from South or Central American countries and have a wide range of variation in ancestry admixture proportions (of EUR, AMR, and AFR alleles). Shaffer (2007) and Fernandez (2013) have reported on associations between West African, EUR, and AMR parental populations, and BMI for adults in the United States.

In this chapter, I briefly explain the literature search strategy and review studies that gave more background on the issues of childhood obesity and disparities in obesity. I then discuss the theoretical framework for my study, evo-devo, which explains how alterations in the mechanisms of embryonic development influence or direct evolutionary changes in any and all stages of the life cycle (Hall, 2012). This chapter also defines the main outcome and predictive variables and fundamental concepts including obesity, BMI, ancestry genetic admixture, as well as epigenetics and environmental and genetic risk factors of obesity. I briefly discuss the economic impact of childhood obesity and disparities in the prevalence of obesity among different races and ethnicities. I go on to discuss the insufficiency of the social constructs of race and

ethnicity in scientific study. Ancestry genetic admixture has been suggested as a surrogate for children's parentally reported race and ethnicity, and so I conclude by reviewing literature on the association between genetic admixture and diseases.

Literature Search Strategy

I searched several electronic databases including PubMed, Google Scholar, the NHGRI-EBI catalog of published GWAS (GWAS Catalog), and Open Thesis, as well as dissertations at Walden. The following key search terms were used: *genetic admixture, epigenetic, childhood obesity, disparity of obesity, risk factors of obesity, Hispanic children obesity, National Health Statistics, consequence of childhood obesity and genetics of obesity, early life obesogenic factors prevalence of childhood obesity, definition of childhood obesity, evolutionary developmental biology,* and *life-history theory*.

The following keyword combinations were used to search for more relevant literature: *Childhood obesity and Hispanics, genetics and childhood obesity, admixture proportion and childhood obesity, admixture proportion and chronic diseases and admixture proportion, Hispanic childhood obesity, and obesity theoretical framework and genetics of complex diseases*. The search was limited to English-language peer-reviewed journals published from 1984 to present. A total of 250 were designated as potentially useful sources; 85 were used in this study. This time frame was chosen because there is no study of admixture proportion and health condition prior to that date.

Evolutionary Developmental Biology

The theoretical framework is a very important aspect of any research process. Theory-driven thinking is crucial for the development of the research topic, and questions. It is like the blueprint of a house. The theoretical framework serves as a structure and support for the study, and it provides a grounding base for literature review as well as the study design, methods, and analysis plan (Grant & Osanloo, 2014). The life-history theory is an evo-devo domain that provides an important conceptual framework to address questions about health and disease (Gluckman, 2011; Hill, 1993).

Evo-devo serves as the theoretical framework for this study to investigate the relationship between growth processes and ancestral background. As Hall (2012) wrote, "Evolutionary Developmental Biology (EvoDevo) is part of biology involved in understanding how alterations in the mechanisms of embryonic development influence or direct evolutionary changes in any and all stages of the life cycle" (p. 184). Calow is zoologist and environmentalist who used the term *evolutionary developmental biology* for the first time in the University of Sheffield in England. Evo-devo theory is concerned with the relationship between changes in embryonic development during single generations and the evolutionary changes between generations. Charles Darwin explained that embryonic development is an important concept to understand in human evolution (Hall, 2012). The evo-devo approach to phenotypic novelty seeks to provide a mechanistic explanation of morphological change (Peterson & Müller, 2016). An appreciation of the fundamental principles of evo-devo provides new insights into major diseases and enables an integrated understanding of human biology and medicine. Public

health professionals are familiar with the physiological basis of disease; an understanding of evo-devo will help them to gain a better understanding of and appreciation for the occurrence of diseases (Gluckman et al., 2011).

Recently, molecular genetics revolutionized the theory of evo-devo by integrating a molecular understanding into evolutionary theory, with operational mechanisms interacting at different levels including gene, cell, tissues, organs, whole organism, and organism–environment. At the gene level, evolutionary developmental mechanisms operate for regulation, networks, interactions, genome size, and epigenetic processes. At the environmental level, evolutionary developmental mechanisms operate for Phenotypic responses to chemicals released by predators, and food supplies (Hall, 2012). Generally, controls on gene regulation and function are considered under the purview of "epigenetics." Epigenetic is a term coined by the British geneticist and embryologist Conrad Waddington for the causal factors that control gene action during development (Hall, 2012).

Epigenetic

According to Hall (2003) epigenetics is "the sum of the genetic and non-genetic factors acting upon cells to control selectively the gene expression that produces increasing phenotypic complexity during development" (p. 492). The genetic aspect of epigenetics has shown that organisms do not start their lives as naked nuclear DNA. They possess DNA in their mitochondria—epigenetic "marks" in their nuclear DNA—and they inherit mRNA and proteins that were produced under the control of their mother's DNA and deposited into the egg cytoplasm (Hall, 2012). Patterns of epigenetic

markers for obese people are different from those of nonobese individuals (Martínez, Milagro, Claycombe, & Schalinske, 2014).

The prevalence of obesity in modern society has two major contributory factors: a historical environmental change, and a genetic predisposition that has its origins in our evolutionary history. From an evolutionary perspective, one of the explanations is that most mutations in the genes that predispose us to obesity are neutral and have been drifting over evolutionary time—so-called "drifty genes" that lead some individuals to be obesity prone and others to be obesity resistant (Speakman, 2016). Obesity results from interactions between environmental and genetic factors. Despite a relatively high heritability of common, nonsyndromic obesity (40-70%), the search for genetic variants contributing to susceptibility has been a challenging task. GWAS have dramatically changed the pace of detection of common genetic susceptibility variants. Recent genome-wide association studies have identified many SNPs associated with adult and childhood BMI (Monnereau¸ 2017). Several genetic variants have been associated with obesity and fat distribution. However, since these variants do not fully explain the heritability of obesity, other forms of variation such as epigenetics marks must be considered.

Epigenetic marks, or "imprinting," affect gene expression without actually changing the DNA sequence. Failures in imprinting are known to cause extreme forms of obesity (e.g., Prader-Willi syndrome) but have also been convincingly associated with susceptibility to obesity. Furthermore, environmental exposures during critical developmental periods can affect the profile of epigenetic marks and result in obesity (Herrera et al., 2011). When methyl groups add to DNA strands, it can affect the activity

of nearby genes. Methylation is controlled by both genetic and environmental factors and altered patterns of DNA methylation are seen in some diseases. It is therefore an ideal biological process to study to determine how race or ethnicity and ancestry contribute to a person's susceptibility to disease (Galanter et al., 2017). Epigenetic modification of the genome through methylation plays a key role in the regulation of diverse cellular processes (Smith, 2013).

Stryjecki, Alyass, and Meyre (2017) have used examples from evolution, heritability, admixture, and monogenic and polygenic studies of obesity to provide explanations for ethnic differences in the prevalence of obesity. Multiethnic studies may provide a better understanding of disparities in obesity to create more targeted and personalized obesity treatments.

Overweight and Obese Children

Obesity is one of the most important causes of chronic disease in the world. Obesity contributes to cardiovascular disease, diabetes mellitus, and other physical and mental chronic disorders. Although severe obesity or monogenic obesity has been more studied in the clinical setting, most chronic conditions are due to moderate or polygenic obesity. Moderate obesity is a multifactorial condition and needs greater attention in the development of a public health strategy for the general population (Grundy, 1998). Childhood obesity is a complicated disease. A child is considered overweight or obese if he or she is above a designated normal weight for his or her age, sex, and height. Unhealthy weight gain in children comprises various factors such as high calorie/low nutrient diets, low levels of physical activity, and genetics (CDC, 2016a).

Childhood obesity is related to several chronic conditions in childhood and adult life including high blood pressure, high cholesterol, cardiovascular disease (CVD), and increased risks of impaired glucose tolerance, type 2 diabetes, asthma, sleep apnea, and joint problems (Bacha, 2016; Cote, Harris, Panagiotopoulos, Sandor, & Devlin, 2013). Furthermore, childhood obesity is associated with anxiety, depression, low self-esteem, and social problems (Morison, 2015). Obese children are more likely to remain obese as adults, and adult obesity is associated with increased risks of heart disease, type 2 diabetes, and cancer (Bass & Eneli, 2015).

Types of Obesity

Monogenic Obesity

Multiple rare forms of obesity are caused by mutations in single genes called monogenic mutations. These mutations in genes control appetite, food intake, and energy homeostasis (Hu, 2008). Obesity is also a characteristic of other genetic syndromes caused by chromosomal abnormalities or mutation, such as Prader–Willi and Bardet-Biedl syndromes. In these syndromes, obesity is present along with mental retardation or reproductive anomalies (Farooqi & O'Rahilly, 2006).

Polygenic or Common Obesity

Although in the 21st century obesity can affect everyone in Westernized and non-Westernized societies, some people tend to be more susceptible. Evidence from animal models, twin studies, and genome-wide association studies of large populations suggests that the variation in human susceptibility to obesity has a genetic component along with obesogenic environmental factors. Despite monogenic obesity, many genes may

influence common or polygenic obesity. Genome-wide association studies scan several genetic markers among thousands of individuals' complete sets of DNA to find gene variations called SNPs such as the fat mass and obesity-associated (*FTO*) gene on chromosome 16. People who carry *FTO* or other obesity-related variants have higher risk of obesity (Loos et al., 2008). Several GWAS among EUR and other racial ethnic groups suggest a partial genetic overlap between obesity loci across various ethnic groups (see Figure 1).

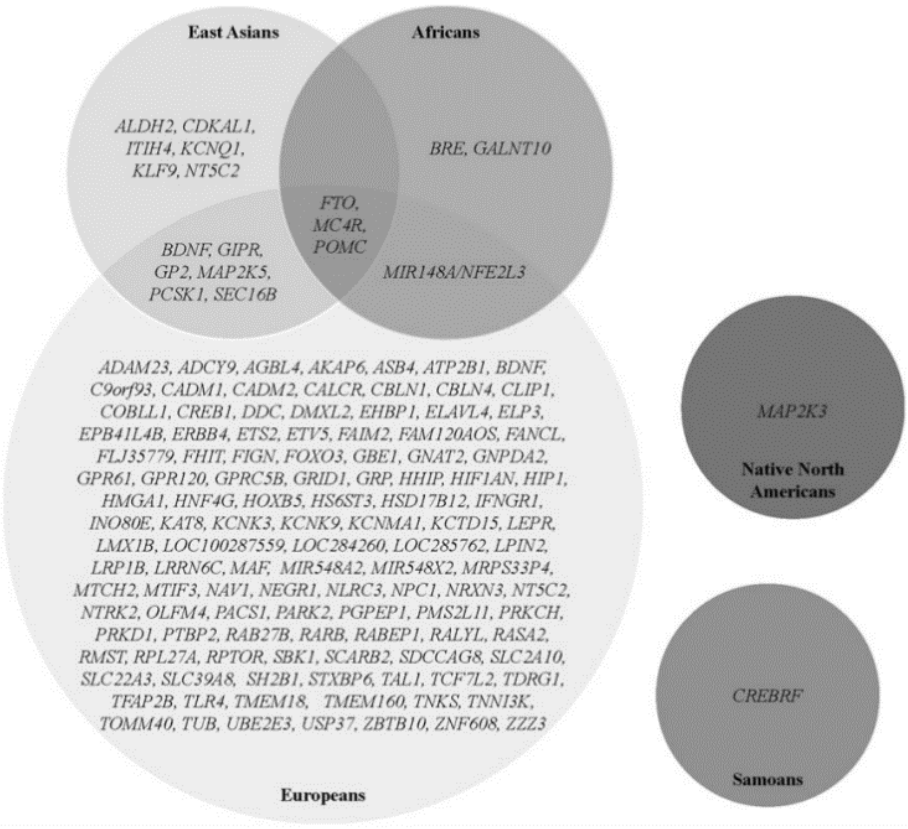

Figure 1. Genetic overlap between obesity loci across various ethnic groups. From "Common Variants Near MC4R Are Associated with Fat Mass, Weight and Risk of Obesity," by R. J. Loos, C. M. Lindgren, S. Li, E. Wheeler, J. H. Zhao, I. Prokopenko, . . . K. L. Mohike, 2008, *Nature Genetics, 40*(6), p. 768.

Prevalence of Childhood Obesity in the United States

The prevalence of childhood obesity has more than tripled in the past five decades (Fryar et al., 2012). The prevalence of obesity among U.S. youth was 17.0% in 2011–2014. Overall, the prevalence of obesity among preschool-aged children (2–5 years; 8.9%) was lower than among school-aged children (6–11 years; 17.5%) and adolescents (12–19 years; 20.5%). The same pattern was seen in both males and females (Ogden, Carroll, Fryar, & Flegal, 2015). However, recent data suggest that the prevalence of

childhood obesity may be decreasing among certain populations. According to the National Heart Lung and Blood Institute (2016), obesity is defined as having excess body fat. Overweight is defined as having excess body weight for a particular height from fat, muscle, bone, water, or a combination of these factors. BMI is a widely used screening tool for measuring whether an individual qualifies as overweight or obese. BMI percentile is preferred for measuring children and young adults (ages 2–20), because this measure considers the fact that they are still growing, and that they could be growing at different rates depending on their age and sex. Health professionals use growth charts to see whether a child's weight falls into a healthy range for the child's height, age, and sex. Children with a BMI at or above the 85th percentile and less than the 95th percentile are considered overweight. Children at or above the 95th percentile are considered obese (CDC, 2016b; see Table 1).

Table 1

BMI Percentile Classifications

Weight Status Category	Percentile Range
Underweight	Less than the 5th percentile
Normal or healthy weight	5th percentile to less than the 85th percentile
Overweight	85th to less than the 95th percentile
Obese	Equal to or greater than the 95th percentile

Body Mass Index Percentile (BMI) and Sex-Specific BMI-for-Age Percentile

BMI is a person's weight in kilograms divided by the square of height in meters (rounded to one decimal place). For children and teens, BMI is age- and sex-specific, and is often referred to as BMI-for-age. According to the CDC (2017), obesity represents a BMI at or above the 95th percentile of the sex-specific CDC BMI-for-age growth charts in children and adolescents aged 2 to 19 years. Overweight was defined by the CDC as possessing a BMI between the 85th and 95th percentiles. There is no recommended definition of obesity in children younger than 2 years; excess weight for children under age 2 was defined as a weight for recumbent length at or above the 95th percentile on the CDC sex-specific weight for recumbent length growth charts (Ogden & Flegal, 2010). A high BMI can be an indicator of high body fatness. Although BMI does not measure body fat directly, research has shown that BMI is correlated with more direct measures of body fat such as skinfold thickness measurements, bioelectrical impedance, densitometry (underwater weighing), dual energy x-ray absorptiometry (DXA), and other methods (Freedman, Horlick, & Berenson, 2013; Garrow & Webster, 1985; Wohlfant-Veje, 2014).

BMI can be considered an alternative to direct measures of body fat. In general, BMI is an inexpensive and easy-to-perform method of screening for weight categories that may lead to health problems.

BMI Percentile

Children's BMI is expressed as a percentile and it can be obtained from either a graph or a percentile calculator. These percentiles express a child's BMI relative to children in the United States who participated in national surveys that were conducted from 1963–1965 and 1988–1994. Because weight and height change during growth and development, as does their relation to body fatness, a child's BMI must be interpreted relative to other children of the same sex and age. The BMI-for-age percentile growth charts are the most commonly used indicator to measure the size and growth patterns of children and teens in the United States. BMI-for-age weight status categories and the corresponding percentiles were based on expert committee recommendations and BMI-for age for 2-year-old boys and girls (24 months) are shown in the following table and graphs (see Tables 2 and 3 and Figures 2 and 3; CDC, 2001).

Table 2

BMI-for-Age for 24-Month-Old Boys

Percentile	BMI Value
3rd	14.52095
5th	14.73732
10th	15.09033
25th	15.74164
50th	16.57503
75th	17.55719
85th	18.16219
90th	18.60948
95th	19.33801
97th	19.85986

Note. Adapted from "Data Table of BMI-for-Age Charts (Males)," by Centers for Disease Control and Prevention (CDC), 2001a. Retrieved from https://www.cdc.gov/growthcharts/html_charts/bmiagerev.htm#males. Copyright 2001 by the CDC.

Table 3

BMI-for-Age for 24-Month-Old Girls

Percentile	BMI Value
3rd	14.14735
5th	14.39787
10th	14.80134
25th	15.52808
50th	16.42340
75th	17.42746
85th	18.01821
90th	18.44139
95th	19.10624
97th	19.56411

Note. Adapted from "Data Table of BMI-for-Age Charts (Females)," by Centers for Disease Control and Prevention (CDC), 2001b. Retrieved from https://www.cdc.gov/growthcharts/html_charts/bmiagerev.htm#females. Copyright 2001 by the CDC.

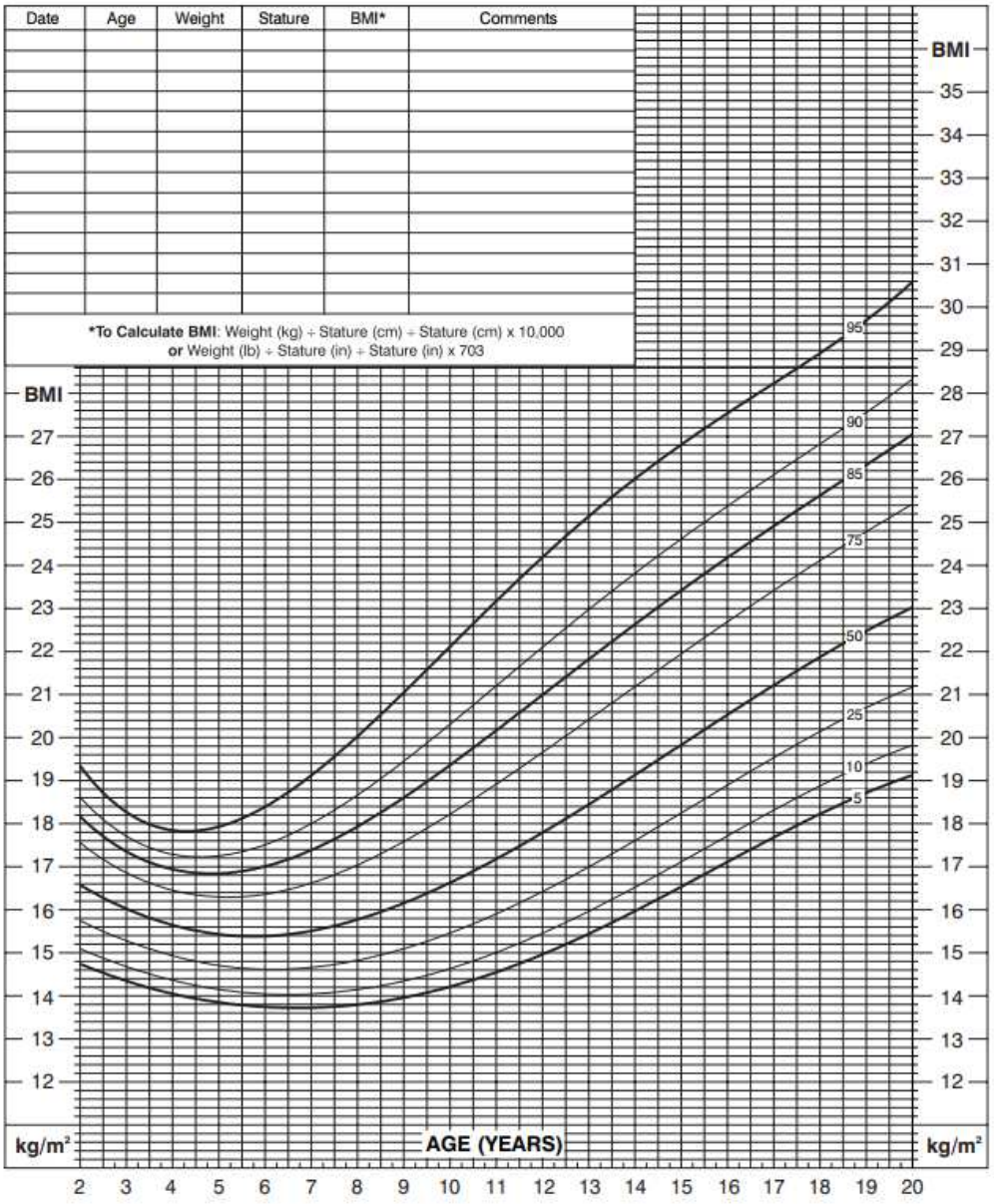

Figure 2. BMI-for-age percentiles, boys 2–20. Adapted from by Centers for Disease Control and Prevention, 2001. Retrieved from https://www.cdc.gov/growthcharts/data/set1clinical/cj41c023.pdf. Copyright 2001 by CDC. Reprinted with permission.

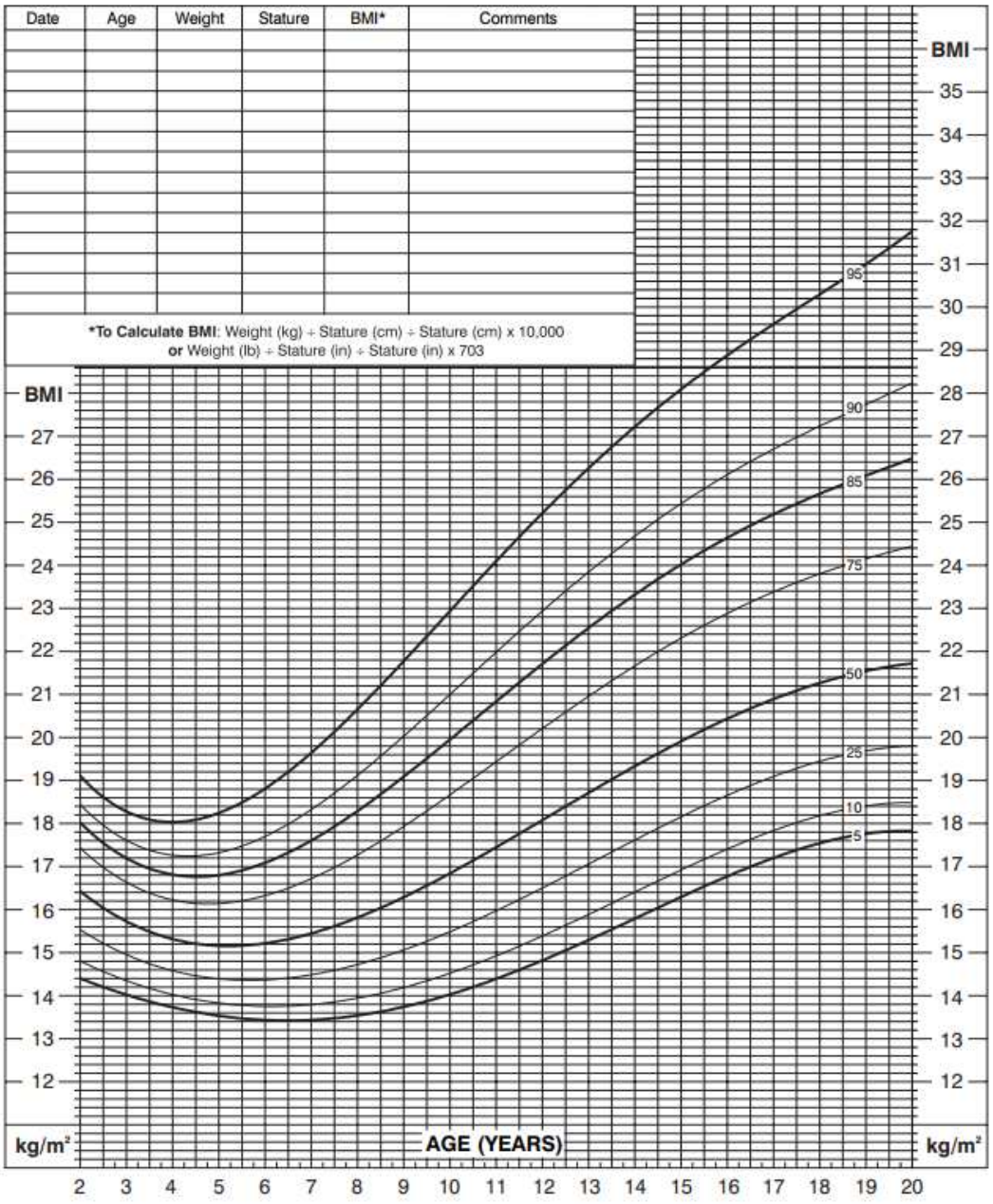

Figure 3. BMI-for-age percentiles, girls 2–20. Adapted from by Centers for Disease Control and Prevention, 2001. Retrieved from https://www.cdc.gov/growthcharts/data/set1clinical/cj41c024.pdf. Copyright 2001 by CDC. Reprinted with permission.

Interpretation of BMI and BMI Percentile

BMI is currently used in the medical profession to screen for weight status and weight-related health problems, and a child with a high BMI for their age and sex should be seen by a health care provider for further assessments such as skinfold thickness measurements, evaluations of diet, physical activity, family history, and other appropriate health screenings. BMI is interpreted differently for children and teens even though it is calculated the same way. Because there are changes in weight and height with age, BMI levels among children and teens need to be expressed relative to other children of the same sex and age.

Economic Impact of Obesity

The high prevalence of childhood obesity is an important burden on U.S. public health and the economy. Furthermore, obesity has a large economic impact on individuals and families (Tremmel, Gerdtham, Nilsson, & Saha, 2017). There are several economic impacts associated with the obesity epidemic including productivity costs (such as disability and premature mortality) and medical costs (which can be as much as 100% higher for an obese individual as compared to a healthy weight adult). Obesity also engenders transportation costs and human capital costs. The overall economic impact of obesity in the United States appears to be substantial (Hammond & Levine, 2010). According to Hammond (2010), the total annual economic costs associated with obesity are about $215 billion. The worldwide economic impact of obesity was estimated to be $2.0 trillion, which is equal to 2.8% of the global gross domestic product (Dobbs et al., 2014).

Risk Factors for Childhood Obesity

Obesity is a global health problem and has increased dramatically in recent decades. Childhood obesity is associated with type 2 diabetes mellitus, cardiovascular disease, some types of cancer, steatohepatitis, and increased risks of premature death in adulthood (Must, Phillips, Naumova, 2012; Williams et al., 2005). Very early life is a critical period, hypothesized to be especially predictive of later obesity risk (Young, 2012). Obesity is a multifactorial disease encompassing genetic, environmental, social, and clinical factors.

Genetics

GWAS have identified multiple genes such as *FTO*, which relate to BMI (Fall & Ingelsson, 2014). Cecil, Tavendale, Watt, Hetherington, and Palmer, (2008) studied 2,726 children for associations with the rs9939609 variant of the *FTO* gene and found strong associations with BMI and weight. Some studies showed that variants in gene *MC4R* are also associated with fat mass, weight, and risk of obesity. In children, these *MC4R* variants may possibly relate to regulation of weight through energy intake and energy expenditure (Loos et al., 2008). Furthermore, a study of 11,653 school children in England found that increased parental BMI is significantly associated with rapid weight gain between the ages of 3 and 5 years (Griffiths, Hawkins, Cole, & Dezateux, 2010).

Social, Environmental, and Behavioral

The fundamental cause of obesity is an energy imbalance between consumed and expended calories. Therefore, obesity is mainly the outcome of feeding or eating behaviors coupled with a sedentary life style. Breastfeeding is found to reduce the risk of

obesity; although controlling for parenting and environmental effects is difficult, breastfeeding is a protective factor according to several studies. Metzger (2010) compared breastfed and nonbreastfed siblings and examined the presence of obesity in adolescence. In sibling pairs in which one was breastfed, and one was not, the breastfed sibling had a lower BMI in adolescence and was less likely to be overweight or obese. Sugar-sweetened beverages are associated with excess weight gain. The largest dietary source of fructose (a lipogenic sugar) provides extra calories; however, these calorie sources may not be well-emphasized compared to solid food such as fast food. These beverages may also replace milk and decrease calcium consumption (Fiorito, Marini, Francis, Smiciklas-Wright, & Birch, 2009).

Screen time is a modifiable risk factor for obesity; most obese children have more than 2 hours of television or computer time a day. African American children have the highest amount of screen time compared to European American and Mexican American children (Vos, 2010). Factors that are associated with increased screen time are lower family income and the presence of a TV in the child's bedroom (He, Harris, Piché, & Beynon, 2009). Examination of cardiorespiratory fitness in school-aged children demonstrated that children with low fitness had a significantly higher risk of being overweight, and had disproportionate increases in weight gain (McGavock, Torrance, McGuire, Wozny, & Lewanczuk, 2009). Family factors are also associated with obesity. The availability of healthy food in the house and familial food preferences influence the foods that children eat. Education, income, poor zip code neighborhoods and lack of access to healthy food and safe outdoor activities, along with some cultural factors are

also associated with increased risks of obesity (Faith & Kral, 2006; Hesketh, Waters, Green, Salmon, & Williams, 2005).

Disparity of Childhood Obesity Among Races and Ethnicities

Childhood obesity is a serious problem in the United States, putting kids at risk for poor health. Despite recent declines in the prevalence of obesity in children aged 2–5 years, obesity amongst all children—particularly Hispanics and African Americans—is still too high. According to the CDC data brief (2017), the prevalence of childhood obesity was 21.9% and 19.5% among Hispanics and non-Hispanic blacks respectively, as compared to 14.7% of non-Hispanic whites. In 2014, the prevalence of obesity was 14.5% for among the Women, Infants, and Children (WIC) participants aged 2–4 years. The prevalence of obesity among these young children was higher for Hispanic children (17.3%) as compared to non-Hispanic White (12.2%; Pan, 2016). The overall prevalence of childhood obesity is higher than the Healthy People 2020 goal of 14.5%, and it is disproportionately higher for Hispanic children (Ogden et al., 2015).

Race and Ethnicity as Social and Cultural Constructs

Self-reported race and ethnicity is used in epidemiological studies to evaluate individuals' origins and understand the roles and interactions between individuals' biology and environment. Although the concepts of race and ethnicity have evolved over time, they are still indistinctly defined terms (Lin & Kelsey, 2000). Usually, participants in the United States specify a race or ethnicity group based on six categories: White, Black, Black Hispanic, White Hispanic, Asian, or other. In the collection of NHANES data, participants reported their self-identified race and ethnicity (Ogden et al., 2014).

Race and ethnicity are related and often used interchangeably; however, race refers to a person's physical appearance—such as skin color and eye color—while ethnicity encompasses cultural heritage, language, social practice, traditions, and geopolitical factors (Mersha, 2015).

Usually one family member declares for the rest, thus preventing detailed analysis of individuals with multiple (and differing) origins. For example, a child of mixed parentage (one black and one white) is typically socially classified as black, even though genetically the child could just as easily be considered white (with a 50/50 genotype). In general, race and ethnicity are problematic and misunderstood terms among scientists (Yudell et al., 2016). The use of self-identified race and ethnicity categories in scientific studies is not sufficient for characterizing the biologic and genetic backgrounds of populations to precisely explore the epidemiology of diseases (Mersha, 2015).

The genetic ancestry of Hispanic individuals widely varies across geographic locations (Bryc et al., 2015). Hispanics who can be of any race are the largest ethnic minority in the United States. They are expected to represent 24% of the U.S. population by 2050 (Lee, 2010). The Hispanic population is genetically diverse, representing a heterogeneous mix of EUR, AFR, and AMR ancestry (Gonzalez et al., 2005). Therefore, a Hispanic individual may self-identify as a single race or as multiple races. There are wide variations across and within Hispanic ethnic groups in terms of genetic, socioeconomic, and cultural factors and geographic origins. Generally, Hispanics have a varied range of ancestry admixture proportions of EUR, AMR, and AFR alleles (Hunninghake et al., 2006). The use of a single Hispanic or Latino ethnic category is

insufficient for characterizing genetic background and disease prevalence (Lara, Akinbami, Flores, & Morgenstern, 2006). For example, Hispanic Americans who are of Mexican origin have a higher proportion of AMR ancestry on average as compared to Hispanic Americans who are of Puerto Rican origin. Puerto Ricans have a higher proportion of AFR ancestry; this may be the reason for the higher prevalence of asthma Puerto Ricans (Chen et al., 2014). Figure 4 compares the proportion of AFR and AMR ancestry for Mexicans and Puerto Ricans (Mersha, 2015).

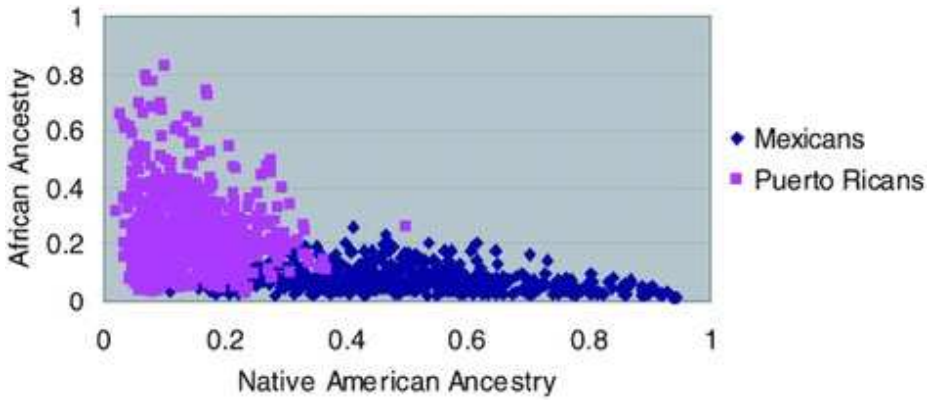

Figure 4. Proportions of African and Native American ancestry for Mexicans and Puerto Ricans.

Ancestry Genetic Admixture, Complex Diseases, and Obesity

The use of parental self-reported race and ethnicity for study of childhood obesity for Hispanic participants can be problematic, because parents may not be fully aware of their own complex ancestry mixture. The complicated genetic structure of Hispanic populations has several important implications for conducting epidemiology studies. Although it is known that prevalence of childhood obesity is higher among Hispanics, there is no population stratification referring to the proportions of their mixed ancestry.

Race, ethnicity, and genetic ancestry have a controversial history in research and practice (Yudell et al., 2016). Race and ethnicity are considered social constructs and cannot capture the heterogeneity within racial and ethnic groups, especially in an admixed population (Borrell, 2005). To account for the heterogeneities, the genetics community has grouped individuals by their genetic ancestry instead of by race and ethnicity (Yudell et al., 2016). Genetic ancestry may better explain the prevalence and disparity of specific medical conditions. However, racial and ethnic categories can usually explain the known risk factors related to the shared cultures, experiences, and exposures such as low socioeconomic status (Nguyen, 2014).

Genetic admixture has been used as a surrogate for biological aspects of race and ethnicity, to investigate the association between ancestry and disease risk. Also, it may help to identify genes that contribute to certain conditions (Fernandez et al., 2013). Stryjecki et al. (2017) have used examples from evolution, heritability, admixture, and monogenic and polygenic studies of obesity to provide explanations for ethnic differences in the prevalence of obesity. Multiethnic studies may provide a better understanding of the disparities of obesity and help to create more targeted and personalized obesity treatments.

Genetic markers, such as ancestry informative markers may provide more accurate information on Hispanics. Ancestry informative markers and newly developed statistical methods are making the genetic estimation of ancestry increasingly more feasible and accurate (Hoggart et al., 2003). According to Skotte et al. (2013), "Admixture occurs when isolated populations begin interbreeding and their offspring

represent a mixture of alleles from different ancestral populations" (p. 693). Genetics admixture is a valuable tool to classify individuals with unknown ancestry and to describe the genetics of different populations (Rosenberg et al., 2002). With next-generation sequencing technologies, it is possible to obtain genetic data for all accessible genetic variations in the genome. For example, using a panel of genetic polymorphisms that present large differences in allelic frequencies between EUR and AFR, it is possible to estimate the degree of EUR and AFR admixture among Hispanics (Ziv et al., 2006).

The 1,000 Genomes Project provides a comprehensive description of common human genetic variation by applying whole-genome sequencing to a diverse set of individuals from multiple populations. This project reconstructed the genomes of 2,504 individuals from 26 populations using a combination of low-coverage whole-genome sequencing, deep exome sequencing, and dense microarray genotyping. They characterized a broad spectrum of genetic variation—over 88 million variants—all phased onto high-quality haplotypes. This resource includes >99% of SNP variants with a frequency of >1% for a variety of ancestries. They describe the distribution of genetic variation across the global sample and discuss the implications for common disease studies (The 1,000 Genomes Project Consortium, 2015). The 1,000 Genomes Project has already elucidated the properties and distribution of common and rare variations, provided insights into the processes that shape genetic diversity, and advanced understandings of disease biology (The 1,000 Genomes Project Consortium, 2012).

The majority of Americans possess an intermixture of EUR, AFR, and AMR ancestries from colonization of New World; therefore, genetic variation of these

populations has established new combinations of genes. Genetic admixture studies may be able to explain differences in populations that cannot be explained by self-reported or socially assigned races or ethnicities. For instance, despite homogenous socioeconomic status, African Americans are 1.5 times more likely to be obese compared to European Americans; this suggests that differences in genetic background may account for ethnic differences in obesity risk (Cheng et al., 2010). Using genome-wide admixture mapping in 15,280 African Americans, Cheng et al. (2010) identified a negative correlation between BMI and the percentage of EUR ancestry, suggesting that the EUR genome may contain either fewer obesity-risk alleles or more obesity-protective genetic factors.

Genome-wide association studies have identified about 100 loci associated to BMI; however, these variants only explain 2.7% of the variation in BMI. Therefore, many other variants remain unidentified due to small sample sizes, genetic heterogeneity, or epigenetic, gene-gene, or more importantly gene-environment interactions (Locke et al., 2015).

Fine-mapping efforts have focused on the *FTO* locus, due to its strong association with obesity-related traits; however, initial attempts to fine-map the causal variant(s) in *FTO* in populations of AFR ancestry have yielded inconsistent results (Lu, 2013). Similar to fine-mapping, admixture mapping is a method for genetic investigation and identifying disease-causing variants in complex diseases like obesity (McKeigue, 2005). Knowing that the prevalence of certain complex diseases such as diabetes and obesity varies with ethnicity, admixture studies scan the genome for regions where the proportion of one ethnicity is significantly different than average. Admixture mapping has a better

statistical power to identify variants with modest effects and has successfully reported associations between risk of obesity and increased BMI in West African and AMR populations (Fernandez et al., 2013). Furthermore, individuals of mixed ancestral background have been found to have an above-average BMI, suggesting that differences in ancestral background may partially explain ethnic differences in the prevalence of obesity (Fernandez et al., 2013).

Epigenetic changes also may explain the missing heritability in obesity. As Suzuki and Bird (2008) wrote, "Epigenetics is defined as changes in gene transcription and expression that do not involve changes to the underlying DNA sequence" (p. xxx). Epigenetic modifications include DNA methylation, histone post translational modifications, and chromatin remodeling or the inheritance of mRNAs that regulate gene expression (Pigeyre, Yazdi, Kaur, & Meyre, 2016). DNA methylation consists of the addition of methyl groups to cytosine residues, and is typically associated with gene silencing (Suzuki & Bird, 2008). Kühnen et al. (2016) reported that methylation within a variably methylated region in proopiomelanocortin is strongly associated with BMI in a multiethnic cohort (Kühnen et al., 2016).

Herrera et al. (2011) argued that common obesity is heritable, but that detecting contributing genetic variants to susceptibility of obesity is very challenging. Furthermore, environmental exposures may affect epigenetic markers and result in obesity. Cardel et al. (2011) evaluated the effect of ancestral genetic background on body composition among African American and EUR children. They reported that, after adjusting for age, height, sex, and socioeconomic status, greater EUR admixture was associated with lower

lean mass. Norden-Krichmar et al. (2014) explored the influence of the degree of AMR admixture on BMI in 846 Native Americans. According to this study, genetic factors may explain some of the variations in obesity among Native Americans. Comuzzie et al. (2012) investigated the genetics of obesity among Hispanic children and found novel genes with unknown function in obesity pathogenesis. Association of AMR ancestry and BMI was investigated and found to be positively correlated in 846 AMR adults; multiple linear regression was used to test the relationship, controlling for socioeconomic and cultural factors (Norden-Krichmar et al., 2014). On the other hand, there is a negative relationship between EUR genetic admixture and the percentage of body fat and BMI among Hispanics and Native Americans (Klimentidis et al., 2009).

The association of genetic ancestry and admixture with other diseases has been studied; for example, Asian admixture was found to be associated with a higher risk for type 2 diabetes among Native Hawaiians (Maskarinec et al., 2016). A case-control study of association of colorectal adenomas and adenocarcinomas with patient ancestral background showed that AFR ancestry was significantly higher in adenomas and cancer cases compared to controls in Columbia (Hernandez-Suarez et al., 2014). Munoz (2016) evaluated the effect of ancestral genetic composition on cardio-metabolic risk factors among Colombian youths. They found that triglyceride levels were associated with the AMR component, and systolic blood pressure was associated with the EUR and AFR components. Furthermore, insulin level resistance was associated with the AFR component. The relationship between genetic admixture and body composition has been studied among Puerto Rican adults living in the United States. EUR ancestry was

associated with lower bone mineral density at the trochanter and femoral neck, and AMR ancestry was associated with lower bone mineral density of the trochanter. AFR ancestry was associated with a higher bone mineral density at the trochanter and femoral neck; however, ancestry was not associated with fat mass, lean mass, or waist circumference (Noel, 2017). Furthermore, EUR ancestry has been identified as a risk factor for multiple sclerosis in Mexicans (Ordoñez et al., 2015).

Genetic ancestry has an association with several chronic diseases. Hispanic American children are heterogeneous and have diverse socioeconomic backgrounds. Hispanic Americans are from any South and Central American country previously under Spanish rule. As discussed previously, prevalence of childhood obesity is higher among Hispanic Americans. The prevalence of childhood obesity within Hispanic Americans may vary by proportion of their genetic ancestry or country of origin. The CDC reports the prevalence of childhood obesity among Hispanics using their reported races and ethnicities; hence, there is a lot of variability in genetic ancestry background of Hispanic Americans that has not been adequately assessed.

Summary

Childhood obesity is a complicated disease; a child is considered overweight or obese if he or she is above a defined normal weight for his or her age, sex, and height. The process of unhealthy weight gain in children is very similar to adults and includes various factors such as a high calorie/low nutrient diet, a low level of physical activity, and genetic factors (CDC, 2016a). The prevalence of childhood obesity has more than tripled in the past five decades (Fryar et al., 2012). The prevalence of obesity among U.S.

youth was 17.0% in 2011–2014. Overall, the prevalence of obesity among preschool-aged children (2–5 years; 8.9%) was lower than among school-aged children (6–11 years; 17.5%) and adolescents (12–19 years; 20.5%). The same pattern was seen in both males and females (Ogden et al., 2015). Children with a BMI at or above the 85th percentile and less than the 95th percentile are considered overweight. Children at or above the 95th percentile are considered obese (CDC, 2017). The high prevalence of childhood obesity is a burden to U.S. public health and economy. Furthermore, obesity has a large economic impact on individuals and families (Tremmel et al., 2017).

According to the CDC (2017), the prevalence of childhood obesity (ages 2–19) was 21.9% and 19.5% among Hispanics and non-Hispanic blacks respectively, as compared to 14.7% of non-Hispanic whites. In 2014, the prevalence of obesity among WIC participants aged 2 to 4 years was 14.5%. The prevalence of obesity among these young children was higher (17.3%) among Hispanic children as compared to non-Hispanic White (12.2%) and non-Hispanic Black children (11.9%; Pan, 2016). The overall prevalence of childhood obesity is higher than the Healthy People 2020 goal of 14.5%, and it is disproportionately higher among Hispanic children (Ogden et al., 2015). Self-reported race and ethnicity is used in epidemiological studies to evaluate individuals' origin and to understand the interactions between an individuals' biology and their environment. Although the concepts of race and ethnicity have evolved over time, they are still indistinctly defined terms (Lin, 2000). Usually, participants in the United States specify a race or ethnicity group based on six categories: White, Black, Black Hispanic, White Hispanic, Asian, or other. In the collection of NHANES data,

participants reported their self-identified race and ethnicity (Ogden et al., 2014). Race and ethnicity are related and often used interchangeably; however, race refers to a person's physical appearance—such as skin color and eye color—while ethnicity encompasses cultural heritage, language, social practice, traditions, and geopolitical factors (Mersha, 2015).

The genetic ancestry of Hispanic individuals varies widely across geographic locations (Bryc et al., 2015). Hispanics of any race are the largest ethnic minority in the United States. They are expected to represent 24% of the U.S. population by 2050 (Lee, 2010). The Hispanic population is genetically diverse, representing a heterogeneous mix of EUR, AFR, and AMR ancestry (Gonzalez Burchard et al., 2005). Therefore, a Hispanic individual may self-identify as a single race or as multiple races. There are wide variations across and within Hispanic ethnic groups in terms of genetic, socioeconomic, or cultural factors and geographic origin. Hispanics have a varied range of ancestry admixture proportions of EUR, AMR, and AFR (Hunninghake et al., 2006). The use of a single Hispanic or Latino ethnic category is insufficient for characterizing genetic background and disease prevalence associated with Hispanics or Latinos (Lara et al., 2006). Genetic admixture has been used as a surrogate for the biological aspect of race and ethnicity in order to investigate association between ancestry and risk of diseases. Also, it may help to identify genes that contribute to or influence the development of certain conditions (Fernandez et al., 2013). Genetic admixture studies may be able to explain differences in populations and certain races or ethnicities that cannot be explained by self-reported or socially assigned races or ethnicities or the environmental and cultural

aspects of races and ethnicities. Although some studies have evaluated the relationship between genetic admixture and childhood obesity, no study has investigated the association of genetic admixture proportions and early childhood obesity among Hispanic American children. I intended to study the influence and contribution of genetic ancestry to variation of prevalence and severity of obesity for Hispanic American children.

In Chapter 3 I present the research questions, hypotheses, variables, research methods, and statistical analyses. I also explain the dependent and independent variables, statistical tests and the process of IRB approval.

Chapter 3: Research Method

Childhood obesity is a major public health concern in the United States and it is higher among ethnic minorities. According to Ogden et al. (2014), 29.8% and 16.7% of Hispanic American preschool children were overweight and obese, respectively; this is twice the national average for this age group. Obesity is a complex disease and many studies have contributed to understanding of the major risk factors of obesity, such as dietary intake and physical activity. Furthermore, epidemiologic studies argue that intrauterine exposure to smoking and higher-than-recommended weight gain during pregnancy may increase glucose levels and result in early onset obesity; in addition, too much screen time and short sleep durations may increase risks for obesity (Taveras et al., 2008).

Racial/ethnic disparities in the rates of childhood obesity among American children are partially explained by socioeconomic and cultural factors; however, disparities within ethnicities are not well investigated. Hispanic American children's parents are originally from South or Central American countries with a wide variety of ancestry admixture proportions of EUR, AMR, and AFR. An association between West African, EUR, and AMR ancestry and adult BMI has been reported in the United States (Fernandez et al., 2013; Shaffer et al., 2007). Cardel et al. (2011) suggested that genetic background may contribute to total body fat accumulation in children. Nevertheless, the association of genetic admixture proportions and obesity in Hispanic American children has not been well investigated. Ancestral genetic background may explain some differences in the prevalence and severity of obesity among Hispanic American children

with similar obesogenic factors and socioeconomic status (Fernandez et al., 2013; Higgins et al., 2005).

This chapter presents the research questions, hypotheses, variables, research methods, and statistical analyses. I present the dependent and independent variables and identify their measurement level in order to demonstrate the choice of appropriate statistical test. The target population is 2-year-old girls and boys whose mothers were recruited into "The First 1,000 Days of Life and Beyond" study in Northern Virginia. The required sample size was calculated using *G**power, and the process of accessing archived data is explained. Then, I describe the study and examine the potential threats to validity. I also discussed ethical procedures, present the IRB approval number for this study, and describe data processing and privacy procedures.

Research Design and Rationale

Study Variables

- Dependent variable: BMI.
- Independent variables: genetic admixture groups including EUR, AFR, and AMR.
- Confounding variables: Sex, birth weight, juice consumption frequency at age 2, maternal education, maternal BMI, maternal age, and maternal region of birth.

Research Design and Connection to the Research Questions

The purpose of this cross-sectional study was to investigate the contribution of children's genetic admixture proportion—including EUR, AMR and AFR—to childhood obesity through the use of BMI among 2-year-old Hispanic American children, adjusted for major obesogenic risk factors of early life such as juice consumption, and

socioeconomic risk factors. The association between genetic admixture data (as independent variables) and children's BMI (as dependent variable) was tested using a simple linear regression. In order to identify significant clinical and social risk factors associated to BMI, a backward stepwise multiple linear regression was conducted to test the contribution of sex, birth weight, juice consumption, mother's age, mother's BMI, mother's education, and mother's region of birth on variability of BMI among Hispanic American children.

Population

The target population were 2-year-old girls and boys whose mothers were recruited to "The First 1,000 Days of Life and Beyond" study in their second or third trimester of pregnancy, and whose parentally reported ethnicity was "Hispanic or Latino." The data for this study was obtained from the Inova Translational Medicine Institute. Inova's "The First 1,000 Days of Life and Beyond" study is a genomic longitudinal cohort based in the Inova Health System in Washington, DC, metropolitan area. About 700 2-year-old children had parentally completed surveys, and about 35% of them were reported as "Hispanic or Latino." The parents' country of birth included 20 South or Central American countries as well as the United States.

Sampling and Sampling Procedures

The archival data was used to identify all 2-year-old Hispanic American children whose mothers are recruited to the "First 1,000 Days of Life and Beyond" study in the prenatal phase of life, and who have available anthropometrics and genomic data as well as demographic, dietary, and clinical data. Children who have medical conditions such as

Hypothyroidism were excluded. Furthermore, children with a physiologically impossible reported height and weight, they were excluded.

Sample Size and Power Analysis

A larger sample size increases the power of statistical tests by collecting more information. Sufficient sample size is essential to test the hypotheses effectively (Kim & Seo, 2013). It is crucial to understand how different research designs require different methods for sample size calculation. In general, sample size depends on three factors: alpha level, effect size, and power level (McCrum-Gardner, 2010). The standard alpha level in the sciences is 0.5; this level is required for the researcher to claim that their discovery is real. By comparing calculated p-value against the set alpha level, we determine whether the observed data are statistically, significantly different from the null hypothesis and ensure that the researcher does not accidentally reject the null hypothesis when it is in fact true (type I error). However, when a difference is statistically significant, it does not necessarily mean that is it important for a conclusion. Therefore, researchers need to calculate the effect size; effect size is usually calculated by taking the difference between the 2 groups (e.g., the mean of case minus the mean of the control group) and dividing it by the standard deviation of one of the groups. Effect size can be calculated after data collection; the common practice is to use a value of 0.5, as it indicates a moderate to large difference (Cohen, 1988). Another important factor for sample size analysis is power level; this helps to avoid a type II error, which refers to a failure to find a statistically significant difference when it actually exists. Power refers to the probability that a test is able to find a statistically significant difference when there is

a real difference. It is generally accepted that power should be .8 or greater; that is, researchers should have an 80% or greater chance of finding a statistically significant difference when there is one (Zint, 2018).

A-priori power analysis was conducted using *G**power for linear regression and a post-hoc power analysis was conducted to assess if the statistical test was able to reject an incorrect null hypothesis (Faul, Erdfelder, Lang, & Buchner, 2007). In *G**power and under test family, I selected *F* tests, and under Statistical test I selected "Linear multiple regression: Fixed model, R^2 increase." Under "Type of Power Analysis," I chose "A priori: Compute required sample size-given α, power and effect size." The minimum required sample size for seven predictors including one admixture proportion and six confounding variables was 103 (see Figure 5). About 150 children were included in this study due to availability of genomic, anthropometrics, social, and clinical data; I calculated effect size given hypothetical sample size using *G**power. The calculated effect size was 0.1 or smaller (Cohen, 1988; see Figure 6).

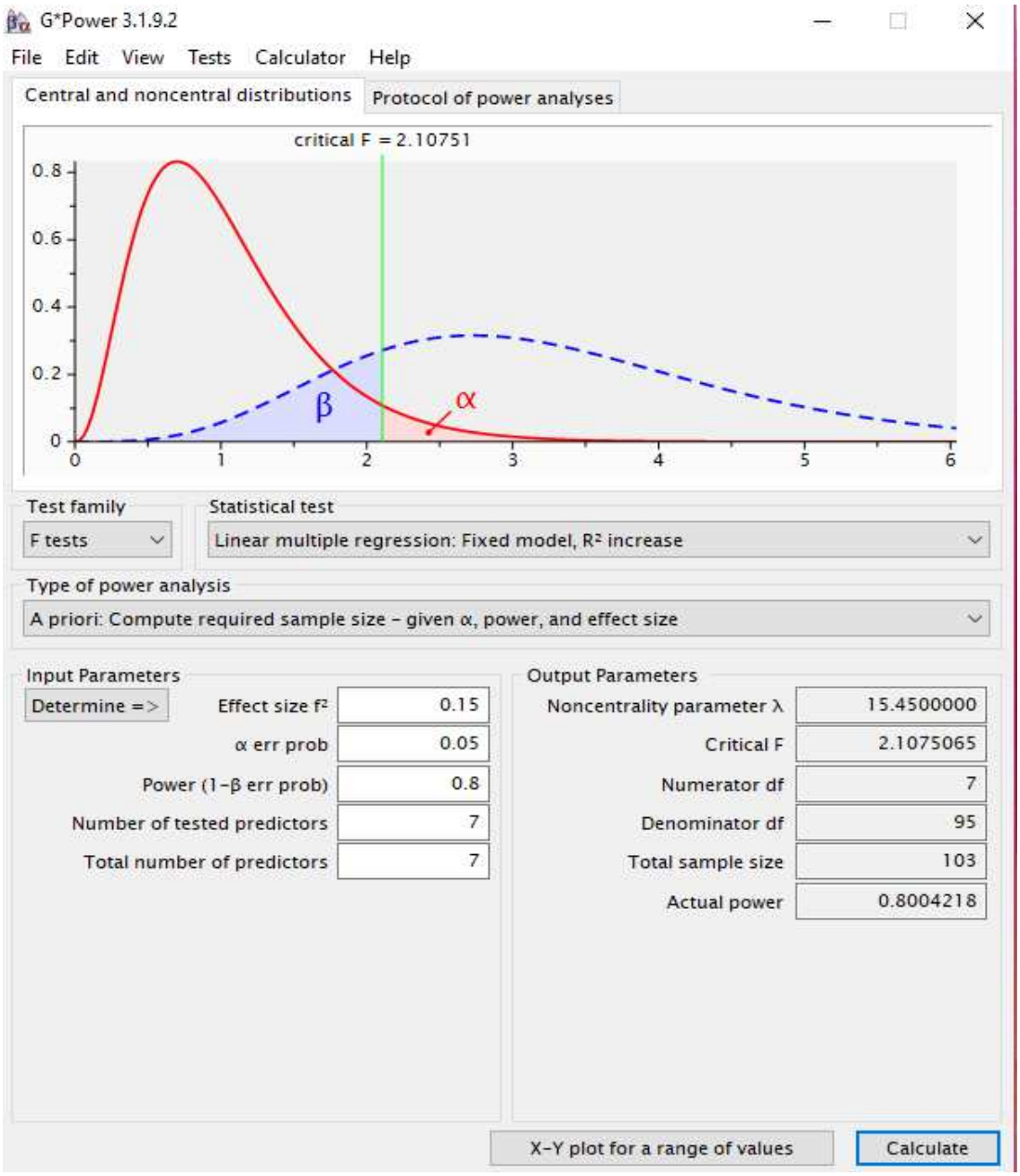

Figure 5. Minimum required sample size.

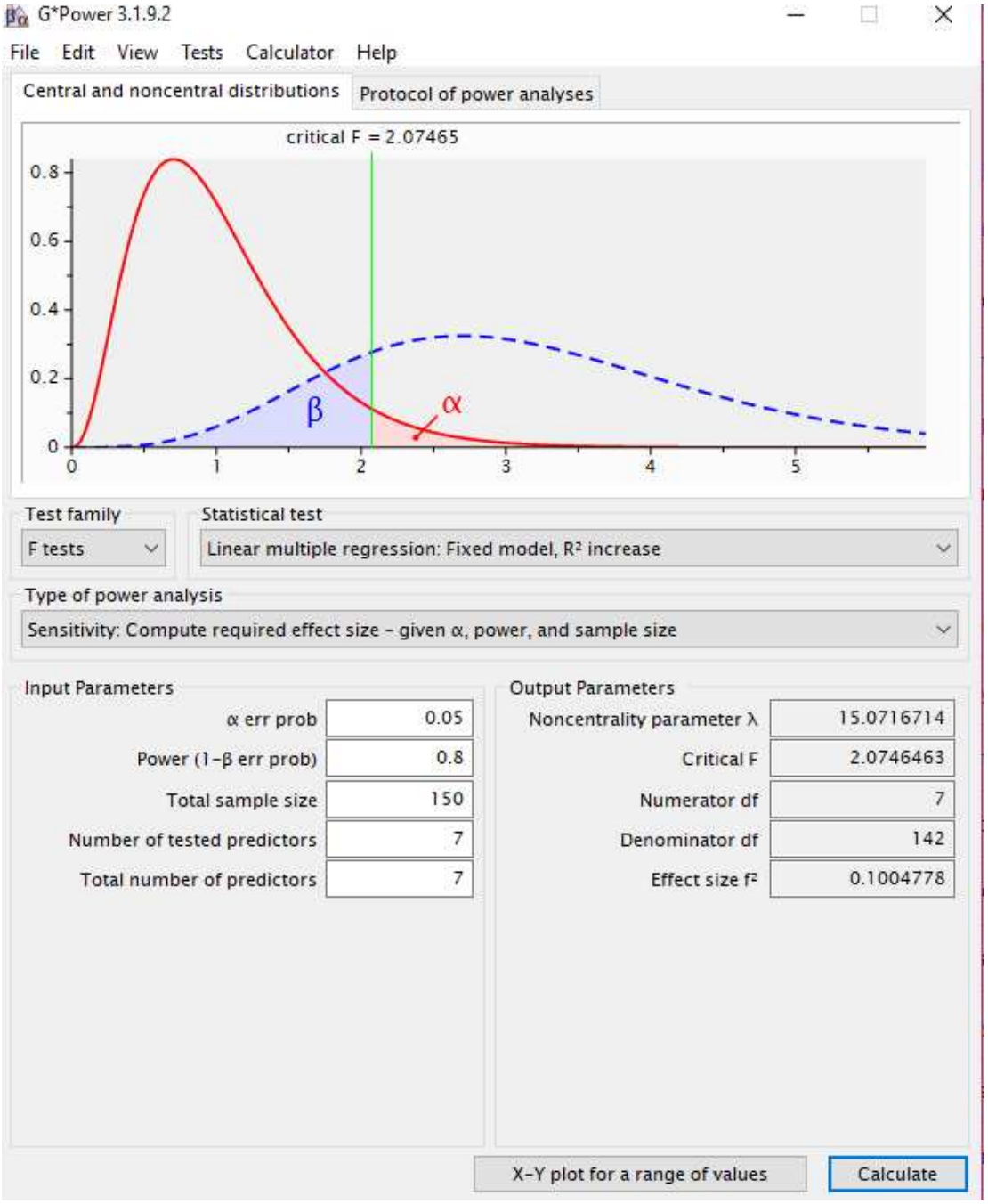

Figure 6. Calculated effect size.

Using Archival Data & Access Procedure

"The First 1,000 Days of Life and Beyond" is an ongoing, family trio, longitudinal cohort study, based in the Inova Health System in Falls Church, Virginia. From 2012 to the present, women have been recruited during pregnancy to participate in the study. Inclusion criteria include being >18 years of age, fluent in English or Spanish, willing to have a biological specimen used for whole genome sequencing (WGS), and have their partner (biological father of fetus) participate in the study. The family then receives surveys every 6 months after delivery until the child is 4-year-old, and then every year until the child is 18. This study was designed to identify genomic, clinical, and environmental risk factors that may enhance our understanding of adverse health outcomes such as premature birth, asthma, obesity, and developmental disorders. IRB approval was obtained for "The First 1,000 Days of Life and Beyond" (WIRB#20120204, Inova IRB#15-1804; Hazrati et al., 2016). To date, about 3,500 families have been recruited to the study; however, only 1400 have WGS. Of those who had WGS performed, about 700 had parentally completed 2-year surveys, and about 35% of them were of Hispanic origin (from South and Central America).

I explained my research question and hypothesis to the principle investigator of "The First 1,000 Days of Life and Beyond" study, and then I reviewed background knowledge and identified the literature gap that helped to formulate my research question. Then, I presented my prospectus and he agreed with my research question and methodology and approved my data use request. I worked with the organization's clinical research manager to get data use permission from institutional IRB. I have been granted

access to the data upon approval of my proposal. I also have access to data collection forms, and data dictionaries.

Instrumentation and Operationalization of Constructs

Admixture Proportion

The 1,000 Genomes Project has described common human genetic variation using WGS of 2,504 individuals from 26 populations in AFR, EAS, EUR, SAS, and AMR using a combination of low-coverage whole-genome sequencing, deep exome sequencing, and dense microarray genotyping (see Figure 7 and Figure 8). They characterized a broad spectrum of genetic variation, over 88 million variants in total.

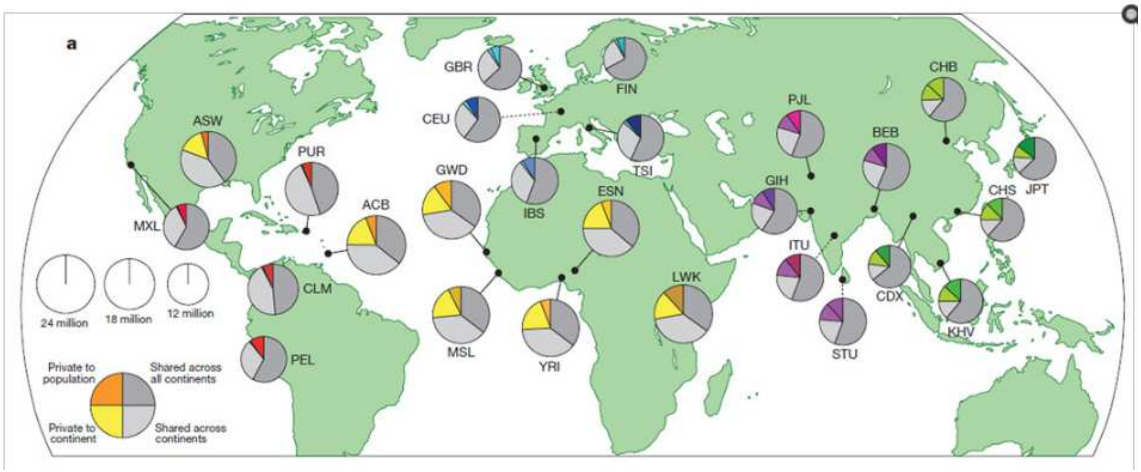

Figure 7. Polymorphic variants within sampled populations.

The area of each pie is proportional to the number of polymorphisms within a population. Pies are divided into four slices, representing variants private to a population (darker color unique to population), private to a continental area (lighter color shared across continental group), shared across continental areas (light grey), and shared across

all continents (dark grey). Dashed lines indicate populations sampled outside of their ancestral continental region (The 1,000 Genomes Project Consortium, 2015).

Population		Code	Population Color	Continental Group Color	Analysis Panel	Phase 1	Phase 3
African ancestry							
Esan in Nigeria	Esan	ESN			AFR		99
Gambian in Western Division, Mandinka	Gambian	GWD			AFR		113
Luhya in Webuye, Kenya	Luhya	LWK			AFR	97	99
Mende in Sierra Leone	Mende	MSL			AFR		85
Yoruba in Ibadan, Nigeria	Yoruba	YRI			AFR	88	108
African Caribbean in Barbados	Barbadian	ACB			AFR/AMR		96
People with African Ancestry in Southwest USA	African-American SW	ASW			AFR/AMR	61	61
Americas							
Colombians in Medellin, Colombia	Colombian	CLM			AMR	60	94
People with Mexican Ancestry in Los Angeles, CA, USA	Mexican-American	MXL			AMR	66	64
Peruvians in Lima, Peru	Peruvian	PEL			AMR		85
Puerto Ricans in Puerto Rico	Puerto Rican	PUR			AMR	55	104
East Asian ancestry							
Chinese Dai in Xishuangbanna, China	Dai Chinese	CDX			EAS		93
Han Chinese in Beijing, China	Han Chinese	CHB			EAS	97	103
Southern Han Chinese	Southern Han Chinese	CHS			EAS	100	105
Japanese in Tokyo, Japan	Japanese	JPT			EAS	89	104
Kinh in Ho Chi Minh City, Vietnam	Kinh Vietnamese	KHV			EAS		99
European ancestry							
Utah residents (CEPH) with Northern and Western European ancestry	CEPH	CEU			EUR	85	99
British in England and Scotland	British	GBR			EUR	89	91
Finnish in Finland	Finnish	FIN			EUR	93	99
Iberian Populations in Spain	Spanish	IBS			EUR	14	107
Toscani in Italia	Tuscan	TSI			EUR	98	107
South Asian ancestry							
Bengali in Bangladesh	Bengali	BEB			SAS		86
Gujarati Indians in Houston, TX, USA	Gujarati	GIH			SAS		103
Indian Telugu in the UK	Telugu	ITU			SAS		102
Punjabi in Lahore, Pakistan	Punjabi	PJL			SAS		96
Sri Lankan Tamil in the UK	Tamil	STU			SAS		102
Total						**1092**	**2504**

Figure 8. Ancestry and kinship toolkit (AKT).

This is a statistical genetics tool for analyzing large cohorts of whole-genome sequenced samples. It can rapidly detect related samples, characterize sample ancestry, calculate correlation between variants, check Mendel consistency and perform data clustering. AKT brings together the functionality of many state-of-the-art methods, with a focus on speed and a unified interface. They believe it will be an invaluable tool for the curation of large WGS datasets (Arthur et al., 2016).

Food Frequency Questions

"The First 1,000 Days of Life and Beyond" food frequency questions were adopted and modified from Behavioral Risk Factor Surveillance System (BRFSS) 2012 Questionnaire (CDC, 2012).

Reported Weights and Heights

Questions were designed and developed by the Inova Translational Medicine Institute Research Team in collaboration with Inova Children's Hospital.

Maternal Education Household Income

Questions were adopted and modified from BRFSS 2012.

Birth Weight and Sex

Birth weight, sex and other birth data were abstracted from the hospital's electronic health records.

Parental Country of Birth

Questions were designed and developed by the Inova Translational Medicine Institute Research Team.

Operationalization of the Variables

Dependent Variable

BMI: BMI is a continuous variable, calculated by dividing parentally reported weight in pounds (lb.) by parentally reported height in inches (in) squared and multiplying by a conversion factor of 703. For instance, if child's weight is 27 lbs. and height is 36 in, then we calculate the BMI as = (27/ (36*36))*703 = 14.64 (CDC, 2014).

Other Terms Used in This Study

Obese: For 2–19-year-old children and adolescents, obesity is defined as having a BMI at or above the 95th percentile of the sex-specific BMI-for CDC Age growth chart (CDC, 2016b)

Overweight: For 2–19-year-old children and adolescents "overweight" is defined as BMI at or above the 85th percentile of the sex-specific BMI-for CDC Age growth chart (CDC, 2016b).

Independent Variables

Admixture: "Admixture occurs when isolated populations begin interbreeding and their offspring represent a mixture of alleles from different ancestral populations." The genetic variation of children in the study were estimated by AKT (Arthur et al., 2016) using 17,535 reliable and common SNPs by projecting the samples into the 1,000 genomes' (Auton et al., 2015) PCs, followed by assigning the PCs to admixture proportions. The 1,000 Genomes Project has described common human genetic variation using WGS of 2,504 individuals from 26 populations in AFR, EAS, EUR, SAS, and the AMR.

Admixture proportions including AFR, AMR, EUR, EAS, and SAS are continuous variables. The sum of the child's admixture proportions should be equal to one; for example: AFR = 0.05, AMR = 0.44, EUR = 0.46, EAS = 0.04 and SAS = 0.01.

Confounding Variables

Juice consumption frequency: a continuous variable collected using a dietary recall form at the 24-month survey. If the reported value was per day, the value is multiplied by seven.

Maternal education: a categorical variable collected through a maternal questionnaire in the prenatal phase or at delivery.

Maternal BMI: a continuous variable calculated from mother's prepregnancy weight and height

Birth weight: a continuous variable and its unit is in grams. Birth weight is abstracted from the electronic health system.

Sex: a categorical variable collected as "Male," "Female," or "Unknown," and it is abstracted from electronic health system.

Mothers' country of birth/region of birth: a categorical variable collected through a maternal and paternal questionnaire in the prenatal phase or at delivery. Country of birth was classified to three regions of birth including South America, Central America, and the United States.

Data Analysis Plan

All statistical analyses and visualization of data was performed using SAS 9.4 (SAS Institute Inc., Cary, NC), R 3.1.2 (R Project for Statistical Computing, Vienna, Austria) and Tableau 10.3.

Data Quality Control, Cleaning, and Preparation

By referring to the data collection forms and data dictionary, data was screened for completeness, correctness, and consistency. Duplicate observations were removed; data was in agreement with data collection tools and the data dictionary. Univariate analysis was performed to describe all independent and dependent variables. Assumptions of linear regression were tested. Distribution of continuous variables was tested using Histogram or Box-and-Whiskers Plots to detect outliers and determine the shape of distribution as well as central tendency and dispersion values. Bivariate analysis was conducted to identify whether collinearity or multicollinearity exist between confounding variables. Frequency of categorical variables was generated to assess the accuracy and completeness of collected data. Using the CDC's sex-specific anthropometric charts, extreme and physiologically impossible heights and weights were excluded from analysis. The sum of the five admixture proportion variables should be equal to one; if the sum of EUR, AME, AFR, EAS, and SAS is not equal to one then the observations would have been excluded. Missing dietary values, were imputed using mean substitution.

Research Questions and Hypotheses

RQ1: Is there an association between children's EUR genetic background and BMI among 2-year-old Hispanic American children?

H_{01}: There is no statistically significant association between children's EUR genetic background and BMI among Hispanic American children.

H_{A1}: There is a statistically significant association between children's EUR genetic background and BMI among Hispanic American children.

To examine RQ1, simple linear regression analysis was conducted to assess if EUR genetic background of children influences BMI. If an association is found, then multiple linear regression would have been conducted to determine whether this association remains significant even after controlling for confounding variables.

RQ2- Is there an association between AMR genetic background and BMI among 2-year-old Hispanic American children ?

H_{02}: There is no statistically significant association between children's AMR genetic background and BMI among Hispanic American children.

H_{A2}: There is a statistically significant association between children's AMR genetic background and BMI among Hispanic American children.

To examine RQ2, simple linear regression analysis was conducted to assess whether AMR genetic background of children is associated with BMI. If an association is found, then multiple linear regression would have been conducted to determine whether this association remains even after controlling for confounding variables

RQ3: Is there an association between children's AFR genetic background and BMI among 2-year-old Hispanic American children.

H_{03}: There is no statistically significant association between children's AFR genetic background and BMI among Hispanic American children.

H_{A3}: There is a statistically significant association between children's AFR genetic background and BMI among Hispanic American children.

To examine RQ3, simple linear regression analysis was conducted to determine whether AFR genetic background of children is associated with BMI. If an association is found, then multiple linear regressions would have been conducted to determine whether this association remains significant even after controlling for confounding variables.

Assumptions of Statistical Tests

Linear regression can assess whether predictive variables explain BMI (criterion); however, the test should satisfy the assumptions of linear regression including normality, linearity, homoscedasticity (Ernst & Albers, 2017). Therefore, prior to analysis, the assumptions of the linear regression model were assessed.

Interpretation of Statistical Tests

Linear regression analysis can estimate the association between a continuous predictive variable and the outcome; multiple linear regression provides a way of adjusting for confounding variables that are included in the model. Multiple linear regression models would have been used to assess whether variations in proportion of genetic background predict children's BMI, while controlling for confounding variables. Statistical significance is set at $P < 0.05$ to determine whether a null hypothesis can be accepted or rejected. Beta coefficients will be used to determine the magnitude of prediction for each independent variable. For everyone unit increase in the significant predictor, the dependent variable will increase or decrease by the number of unstandardized beta coefficients (Statistics Solutions, 2013).

Threats to Validity

The study subjects are 2-year-old Hispanic American children who live in Northern Virginia and the Washington, DC, metropolitan area. All the children were born in the Inova Health System, one of the largest obstetrics and gynecology hospitals in Northern Virginia. The children's parents are originally from 20 South or Central American countries, and therefore, the children are a good representation of the 2-year-old Hispanic American population in the area. However, external validity maybe jeopardized since these study participants may not represent all Hispanic American children in the United States due to differences in acculturation from state to state; therefore, we should not generalize these findings. One threat to internal validity would be using parentally reported survey data, which is not always accurate because parents may give subjective answers. In this study, anthropometrics and dietary values have been reported by parents. The prevalence of childhood obesity is underestimated when using parentally reported weight and height measurements (Scholtens et al., 2006).

Ethical Procedures

Secondary data using archived data from "The First 1,000 Days of Life and Beyond" was used in this study. The main study was designed to identify genomic, clinical, and environmental risk factors that may enhance our understanding of adverse health outcomes (such as obesity). IRB approval was obtained from Walden University (Approval No. 07-03-18-042050) and Inova Health System IRB (WIRB#20120204, Inova IRB#15-1804). A data use agreement was received from Inova Translational Medicine Institute. Access was granted to the data, data collection forms, and data

dictionaries after approval of my proposal and completion of IRB requirements. All the requested data were de-identified and did not include any protected health information such as name, address, or date of birth. Data is stored in a Health Insurance Portability and Accountability Act (HIPAA) compliant encrypted flash drive, and it will be destroyed after completion of my PhD.

Summary

Chapter 3 briefly explained the research question, gap in the literature, and study variables, and the measurement level of the variables. This is a retrospective cross-sectional study to investigate the contribution of children's genetic admixture proportions (of EUR, AMR and AFR) to childhood obesity (measured through children's BMI) among 2-year-old Hispanic American children. I discussed the methodology for evaluating research questions and testing the hypotheses. Required sample size was calculated for seven independent variables (including one admixture variable at a time and six confounding variables). Three separate models tested the association of BMI to EUR, AMR and AFR individually. The confounding variables were: (a) frequency of juice consumption at age 2; (b) sex; (c) maternal education; (d) mother's BMI; (e) mother's age; (f) mother's region of birth; and (g) birth weight (Brophy, 2009; Sahoo et al., 2015). The outcome variable is children's calculated BMI. Threats to validity and ethical considerations were explained in this chapter. Descriptive and inferential analyses were conducted using R 3.1.2 and SAS 9.4.

The details of the statistical analysis for this study are explained in Chapter 4, and research questions were evaluated using descriptive and inferential analyses.

Chapter 4: Results

Purpose of the Study

The purpose of this quantitative, cross-sectional, research study was to investigate the contribution of children's genetic admixture proportion—including EUR, AMR and AFR—to childhood obesity through the use of BMI among 2-year-old Hispanic American children. The association between genetic admixture data (as independent variables) and children's BMI (as the dependent variable) were investigated using "The First 1,000 Days of Life and Beyond" study's secondary data.

Research Questions and Hypotheses

RQ1: Is there an association between children's EUR genetic background and BMI among 2-year-old Hispanic American children?

H_{01}: There is no statistically significant association between children's EUR genetic background and BMI among Hispanic American children.

H_{A1}: There is a statistically significant association between children's EUR genetic background and BMI among Hispanic American children.

RQ2: Is there an association between AMR genetic background and BMI among 2-year-old Hispanic American children?

H_{02}: There is no statistically significant association between children's AMR genetic background and BMI among Hispanic American children.

H_{A2}: There is a statistically significant association between children's AMR genetic background and BMI among Hispanic American children.

RQ3: Is there an association between children's AFR genetic background and BMI among 2-year-old Hispanic American children.

H_{03}: There is no statistically significant association between children's AFR genetic background and BMI among Hispanic American children.

H_{A3}: There is a statistically significant association between children's AFR genetic background and BMI among Hispanic American children.

Organization of Chapter 4

This chapter delineates the statistical analysis and study findings in regard to the research questions and the study hypotheses. The chapter also explains data management procedures and describes the study cohort in detail. I present the rationale for statistical analysis as well as a detailed description of calculated and derived variables.

Data Collection, Management, and Quality Control

Archival data for all 2-year-old Hispanic American children whose mothers were recruited into "The First 1,000 Days of Life and Beyond" study in April 2012–January 2016 was used. The de-identified data file included 208 subjects and was provided to me in a CSV format through Inova Outlook e-mail. I excluded four duplicate records, 32 subjects who had missing reported height or weight, and 16 extreme outliers with physiologically impossible height or weight values (see Data Management, below). Two children were excluded due to hypothyroidism. The total final sample size for this study was 154 Hispanic American children.

Prior to the data analyses, I screened data for inconsistency, missing values, and outliers using the data collection forms and data dictionary as data quality metrics. I

screened the data for completeness, correctness, and consistency. Duplicate observations were removed, and univariate analysis was performed for all independent and dependent variables. Distribution of continuous variables was tested using histogram and box-and-whiskers plots to detect outliers and determine the shape of distribution as well as central tendency and dispersion values. Frequency of categorical variables was generated to assess the accuracy and completeness of collected data. The sum of the five admixture proportion variables was calculated for each entry to assure that it added up to 1. Maternal countries of birth were categorized into three regions: South America, Central America, and the United States. BMI was calculated as weight in pounds divided by the square of height in inches multiplied by 703 (rounded to one decimal place). Using the CDC's sex-specific anthropometric BMI-for-age growth charts, extreme and physiologically impossible heights and weights were excluded from analysis, and I categorized BMI to underweight, normal weight, overweight, and obese (Table 4). The average BMI of the children was 17.4 ($SD \pm 2.6$).

Table 4

BMI Percentile Categories by Frequency and Percent

BMI Percentile Categories	Number of Children	Percent
Under Weight	13	8.5
Normal	85	55.2
Overweight	25	16.2
Obese	31	20.1

Results: Descriptive Statistics and Analyses

Descriptive statistics were used to report the main characteristics of the study cohort and obesity risk factors (Table 5). Exact distribution of children's BMI is shown in Figures 9 and 10.

Table 5

Demographic Characteristics

Characteristics	Frequency or Mean	Percentage or ±Standard Deviation
Sex		
Male	72	47.1%
Female	82	52.9%
Birth weight (gram)	3346.6	±504.8
Juice consumption frequency at 24M (per week)	2.0	±2.3
Maternal age (years)	30.5	±5.1
Maternal BMI	26.2	±4.7
Maternal ethnicity (Hispanic or Latino)	155	100%
Maternal education level		
Less than associate	114	74.2%
Associate degree and above	40	25.8%
Maternal country of birth (region)		
U.S.A.	22	14.3%
South America	37	24.0%
Central America	95	61.7%

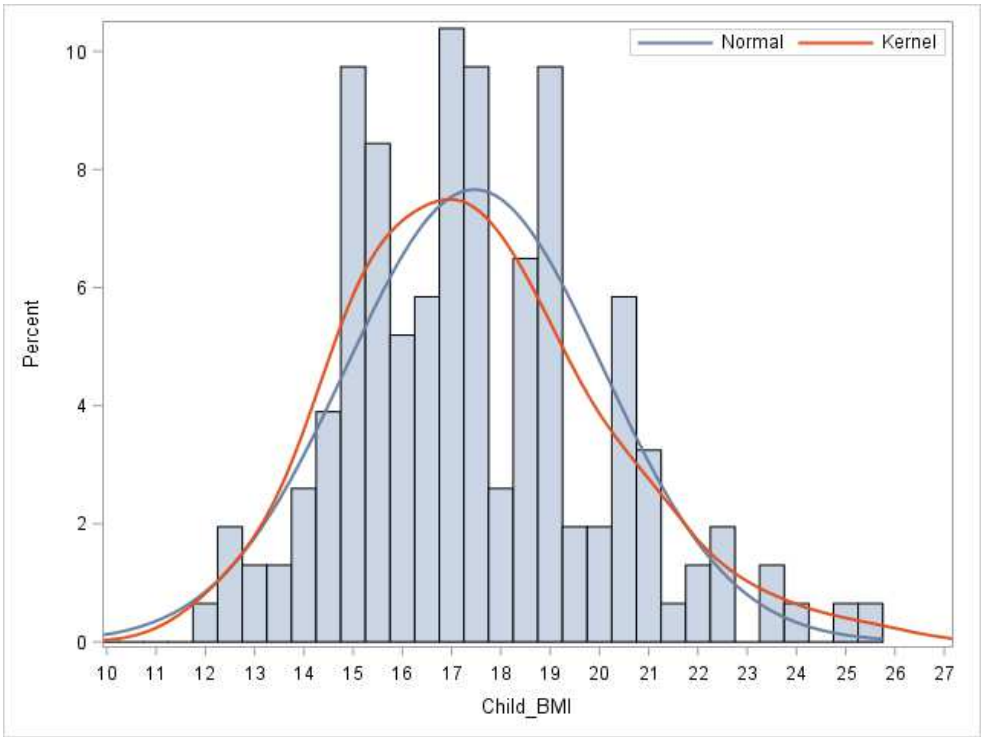

Figure 9. Distribution of children's BMI.

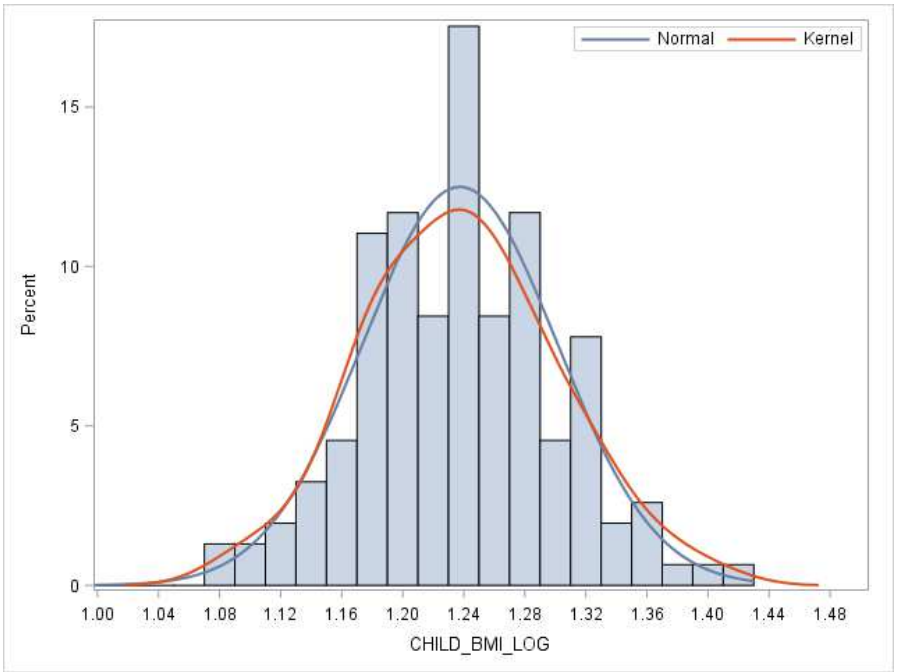

Figure 10. Distribution of children's BMI (log transformed).

The mean (± *SD*) five super group genetic admixture composition of the study cohort was EUR 0.43(*SD* ± 0.23), AMR 0.44(*SD* ± 0.22), AFR 0.08(*SD* ± 0.07), EAS 0.03 (*SD* ± 0.04) and SAS 0.02(*SD* ± 0.03). Distribution of genetic admixture of children is presented in Figure 11 and distribution of each admixture including AMR, EUR, AFR, EAS, and SAS is shown in Figures 12–16.

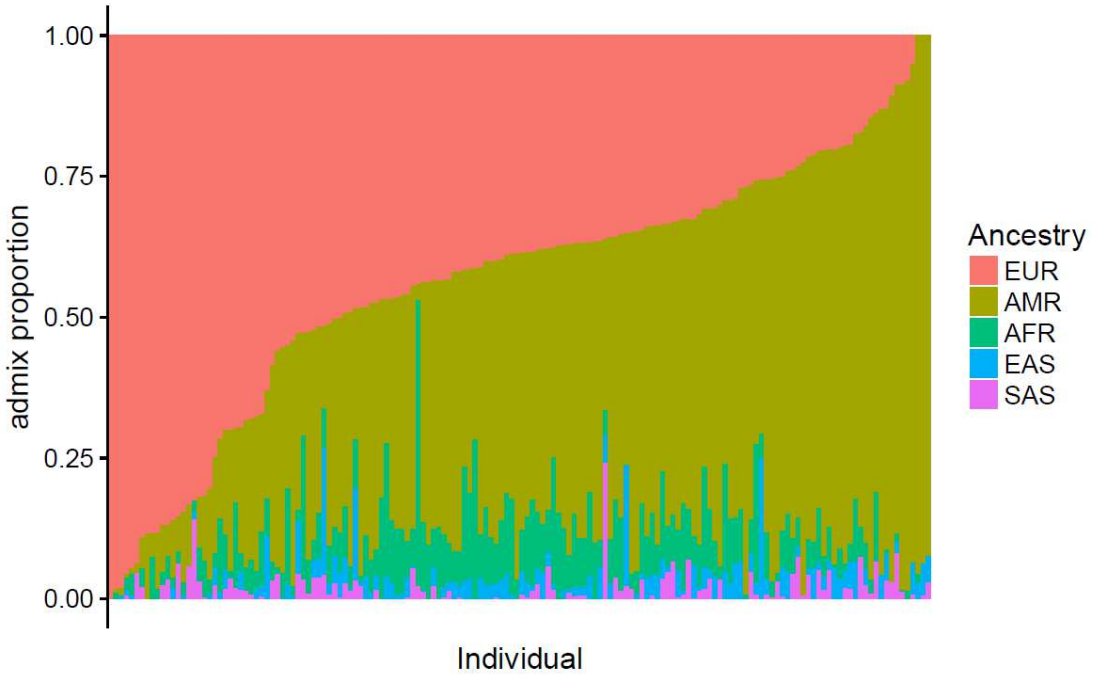

Figure 11. Distribution of five super population genetic admixture.

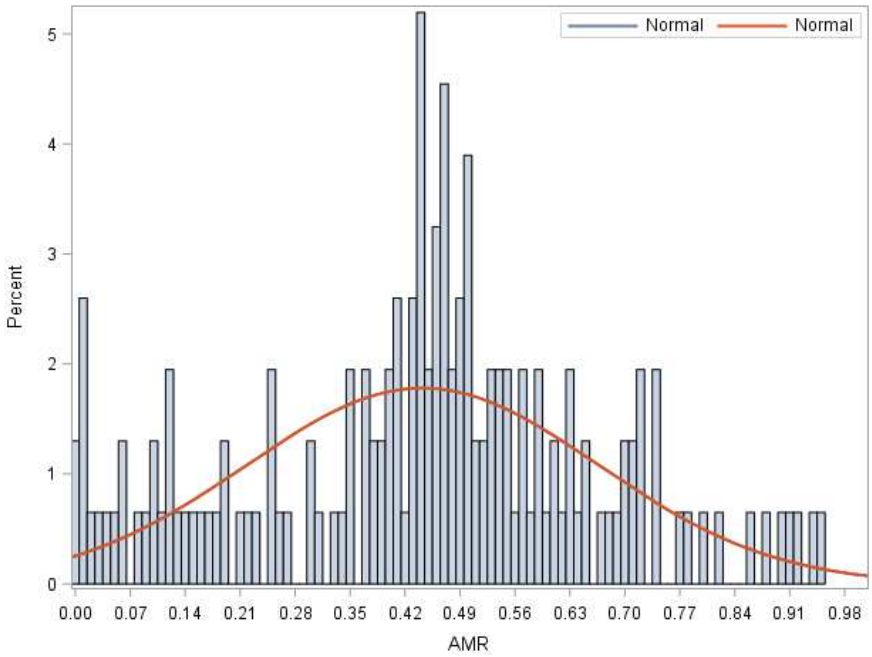

Figure 12. Distribution of American (AMR) ancestry.

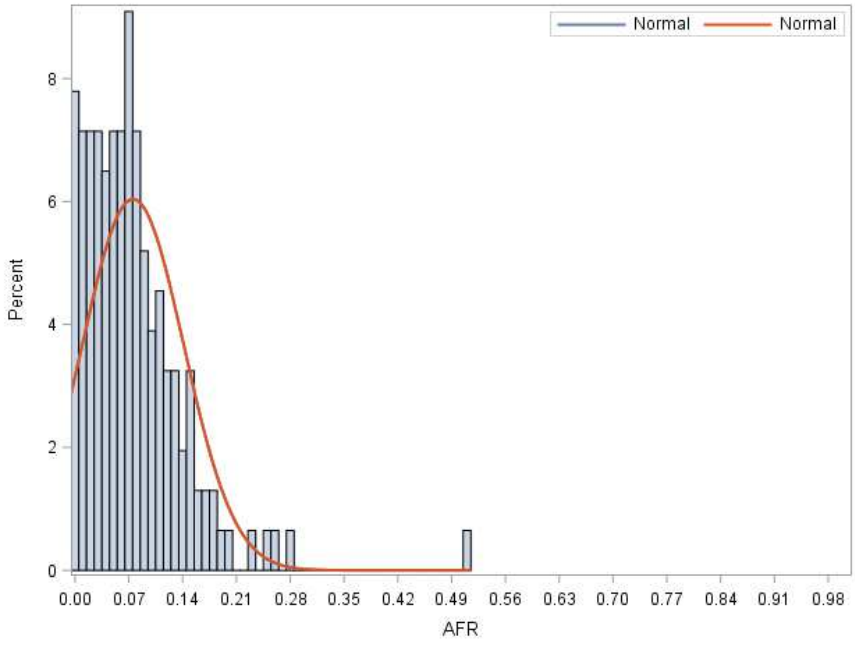

Figure 13. Distribution of African (AFR) ancestry.

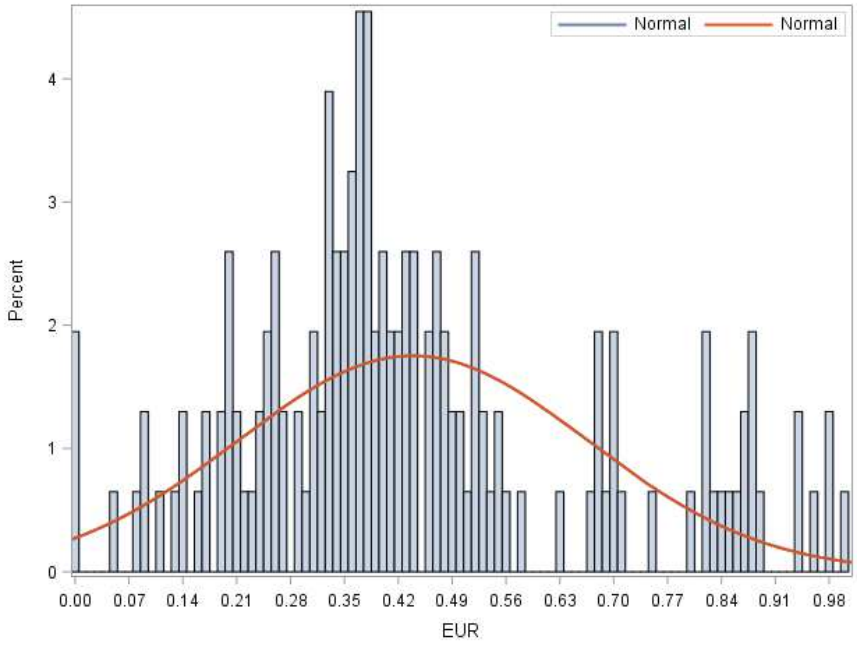

Figure 14. Distribution of European (EUR) ancestry.

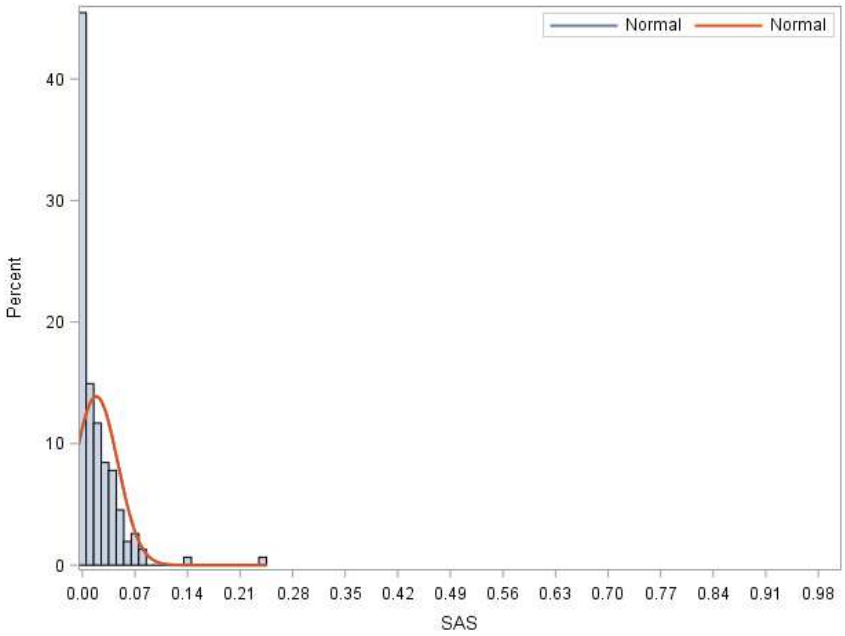

Figure 15. Distribution of South Asian (SAS) ancestry.

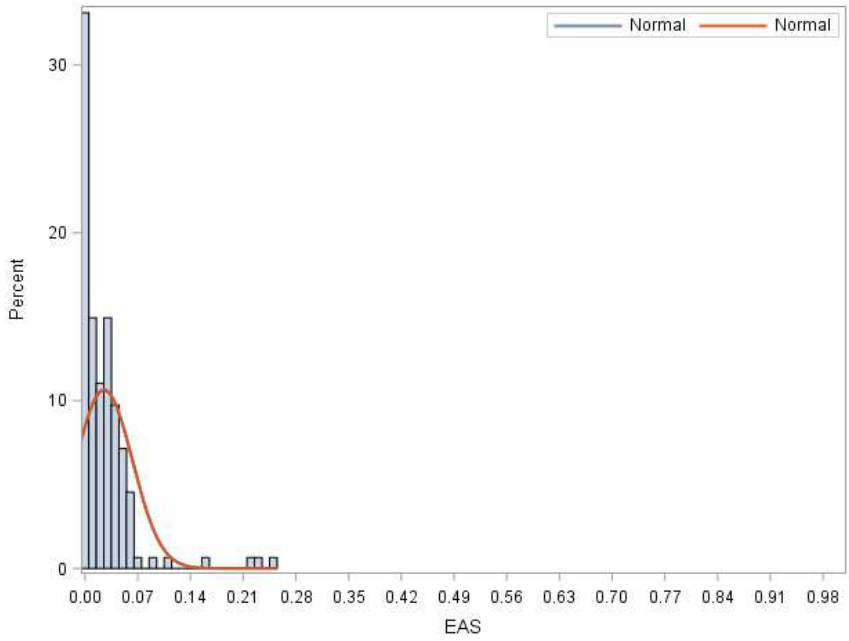

Figure 16. Distribution of East Asian (EAS) ancestry.

Distribution of maternal age (30.5(± 5.1)), maternal BMI (26.2(± 4.7)), child birth weight (3346.6(± 504.8)) and juice consumption frequency at 24 months (2.0(± 2.3)) are displayed in Figures 17–21 respectively.

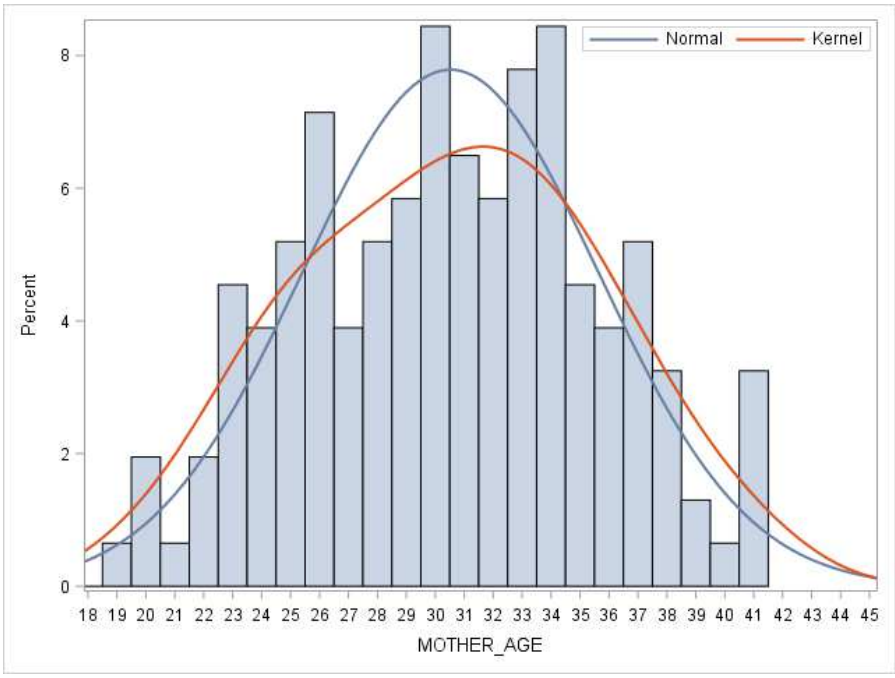

Figure 17. Distribution of maternal age at delivery.

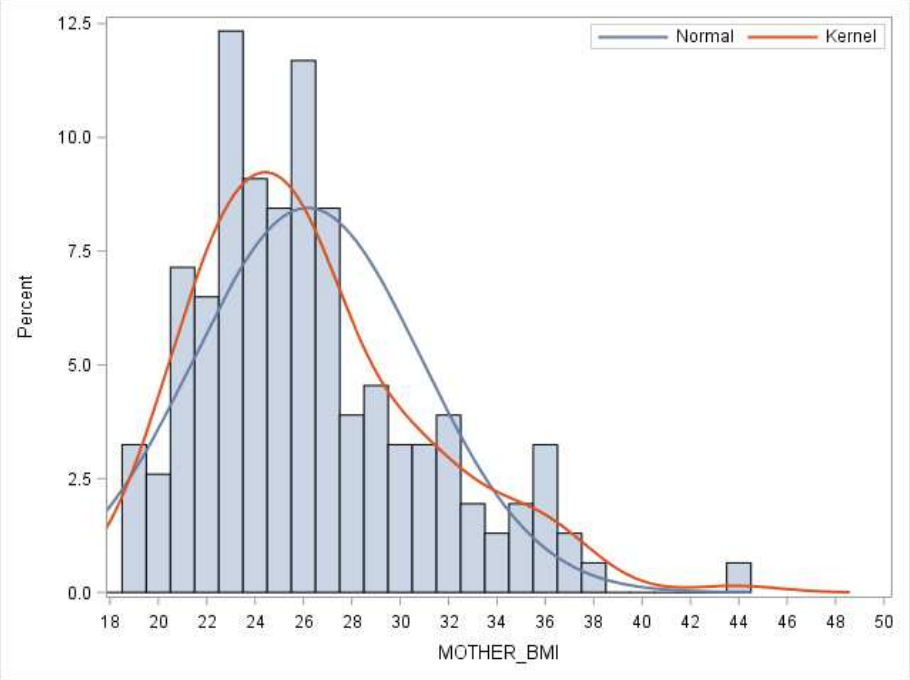

Figure 18. Distribution of maternal prepregnancy body mass index (BMI).

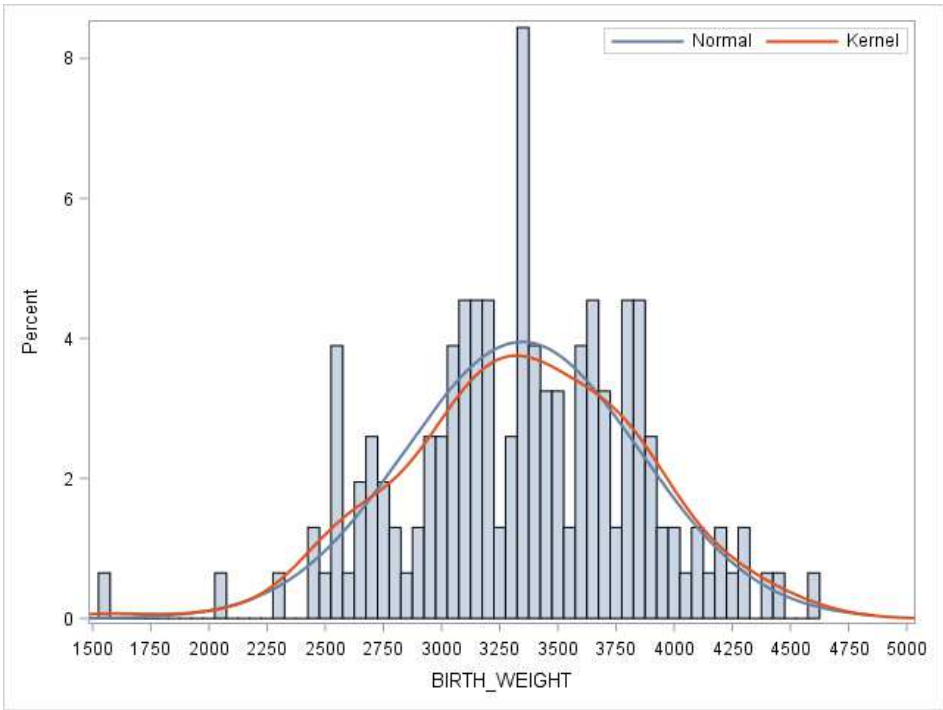

Figure 19. Distribution of maternal child birth weight.

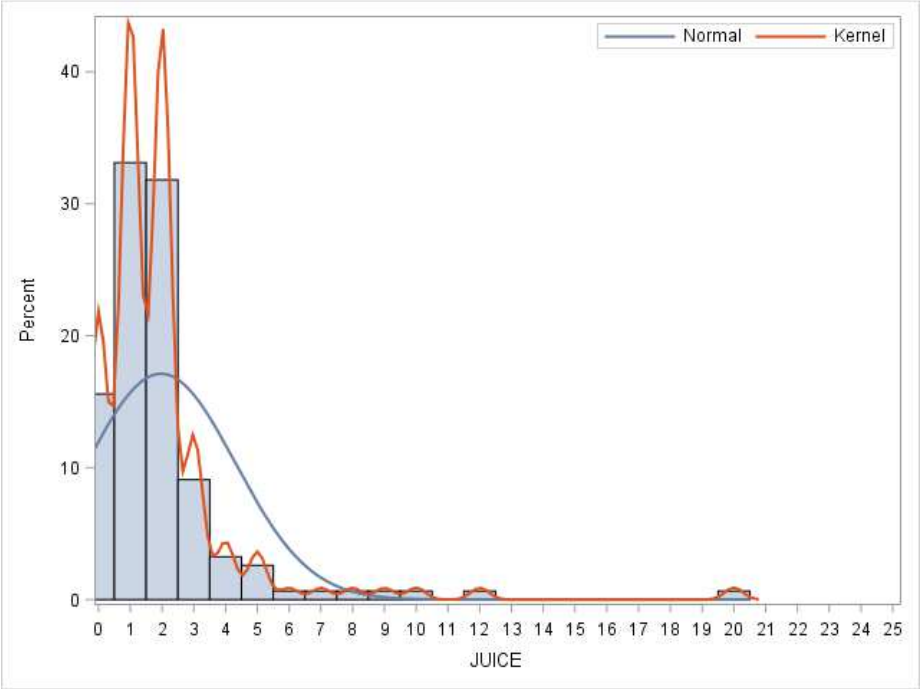

Figure 20. Distribution of juice consumption frequency per week at 24 months.

Inferential Statistical Analyses

Potential Covariates: Evaluating Environmental, Clinical, and Social Risk Factors Associated with BMI

In order to identify significant risk factors associated with BMI, a backward stepwise multiple linear regression was conducted to test the contribution of sex, birth weight, juice consumption, mother's age, mother's BMI, mother's education, and mother's region of birth on variability of BMI among Hispanic American children. Assumptions of multiple linear regression were tested. BMI was normally distributed; however, for regression analysis the normality test should be applied to the residuals rather than the raw values. Due to small sample size, I used the Shapiro-Wilk test to test the assumption of normality. The Shapiro-Wilk test rejected the null hypothesis for

normality ($P = 0.03$); therefore, BMI was log transformed for normality. After log transformation, the Shapiro-Wilk test failed to reject the null hypothesis for normality ($P = 0.2$; see Figures 21 and 22). BMI, birth weight, juice consumption, mother's age, and mother's BMI were measured at the continuous level. Sex, mother's education, and mother's region of birth were dummy coded to numeric values. No collinearity or multicollinearity was detected (Variance inflation factor < 1.2). Significant outliers of BMI were removed in the data cleaning phase. Results from backward stepwise multiple linear regression revealed that only birth weight and maternal education were left in the model due to significance levels of 0.15. Birth weight was positively associated to BMI ($p = 0.02$) and higher maternal education was negatively associated to BMI ($p = 0.03$); the model was significant for the association of BMI and the selected variables (F (2, 151) = 3.6, $p = 0.01$).

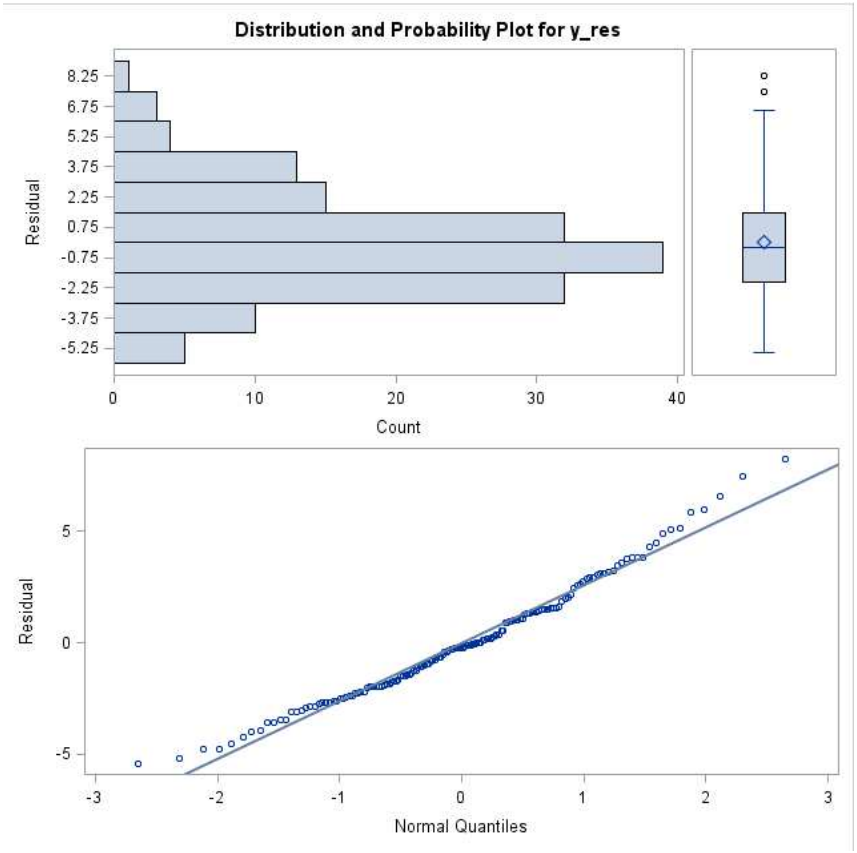

Figure 21. Distribution of BMI residuals.

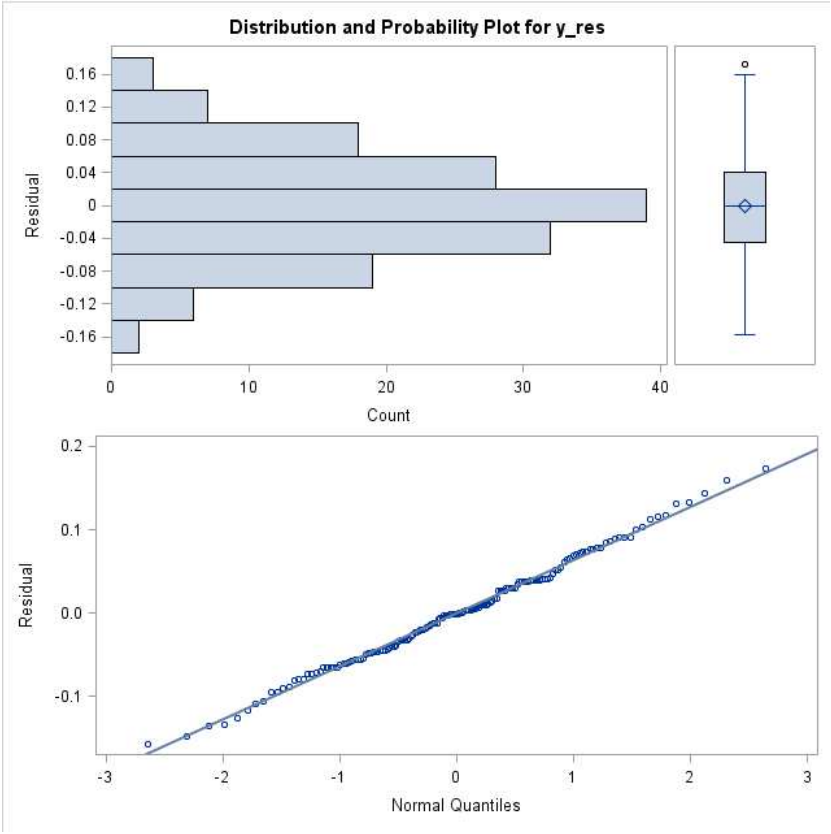

Figure 22. Distribution of log transformed BMI residuals.

Association of Genetic Admixture Proportion to BMI

To examine RQ1, simple linear regression analysis was conducted to assess if EUR genetic background of children is associated to BMI. Assumptions of simple linear regression were tested. BMI and admixture proportion data are measured at the continuous level. There was a linear relationship between BMI and admixture proportion. Significant outliers were removed. BMI was log transformed for normality. A plot of residuals versus predicted values was generated, the residuals variance was around zero indicating that the assumption of homoscedasticity is not violated (Figure 23).

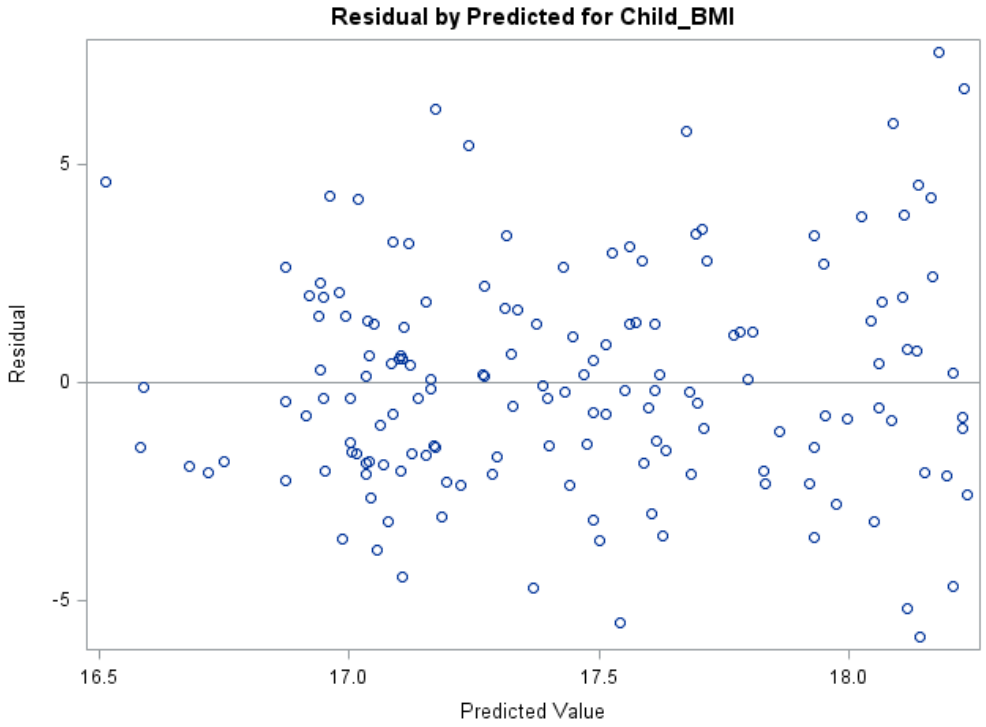

Figure 23. Residual by predicted for children's BMI.

A simple linear regression was calculated to evaluate the relationship between BMI and EUR. No significant association was found between BMI and EUR (F (1, 152) = 0.02, p = .87; see Table 6). To examine RQ2, a simple linear regression was calculated to predict BMI based on AMR. No significant association was found (F (1, 152) = 0.00, p = .97; Table 6). To examine RQ3, a simple linear regression was calculated to predict BMI based on AFR. No significant association was found (F (1, 152) = 0.02, p = .88; Table 6).

Table 6

Linear Regression for BMI Based on Genetic Background

Genetic Background	*B*	*SE B*	ß	*t*	*p*
EUR	0.004	0.02	0.01	0.15	0.87
AMR	0.001	0.02	0.003	0.04	0.97
AFR	0.01	0.08	0.01	0.13	0.88

Furthermore, genetic admixture proportion was plotted and compared among underweight, normal weight, obese and severely obese children (Figure 24).

Figure 24. Comparison of genetic admixture proportion among underweight, normal weight, overweight, and obese children.

Summary

In this cross-sectional quantitative study, analyses were conducted to assess whether there was an association between genetic ancestry background and BMI among 2-year-old Hispanic children living in the DC metropolitan area. The association between genetic admixture composition and BMI were investigated using "The First 1,000 Days of Life and Beyond" study's secondary data. The average BMI of the children was 17.4 ($SD \pm 2.6$), and of 154 children 8.5% were under weight, 55.2% were normal weight, 16.2% were overweight and 20.1% were obese. The distribution and proportion of five super group genetic admixture composition of children were assessed and visualized. Potential clinical, environmental, and social risk factors of childhood obesity were examined using backward stepwise multiple linear regression, only birth weight and maternal education were associated to BMI.

The results for the first research question based on simple linear regression revealed that there is not an association between children's EUR genetic background and BMI among 2-year-old Hispanic American children. The results for the second research question based on simple linear regression revealed that there is not an association between children's AMR genetic background and BMI among 2-year-old Hispanic American children. The results for the third research question based on simple linear regression revealed that there is not an association between children's AFR genetic background and BMI among 2-year-old Hispanic American children.

In the Chapter 5, I summarize the key findings and compare the results with what I reviewed in the peer-reviewed literature described in Chapter 2. I also assimilated the

findings in the context of theory of evolutionary developmental biology. Furthermore, I explained the limitations of the study, implications for social change, and recommendations for future research. The chapter concluded with applicable remarks to finalize this study.

Chapter 5: Discussion, Conclusions, and Recommendations

Estimating genetic admixture has allowed scientists to investigate the relationship between genetic ancestry and diseases in admixed populations (Samet, 1988). Although it is known that the prevalence of childhood obesity is higher among Hispanics, there is no population stratification referring to the proportions of their mixed ancestry. In order to investigate the association between ancestry and disease risk, genetic admixture has been used as a surrogate for biological aspects of race and ethnicity (Yudell et al., 2016). Hispanic American children are the largest ethnic minority in the United States. The Hispanic population is genetically diverse, representing a heterogeneous mix of EUR, AFR, and AMR ancestry (Gonzalez et al., 2005); however, parental self-reported race and ethnicity is not sufficient to capture all the ancestral heterogeneity in Hispanic American children. The complicated genetic structure of Hispanic populations has several important implications for epidemiological studies.

The purpose of this quantitative, cross-sectional research study was to investigate the contribution of children's genetic admixture proportion—including EUR, AMR and AFR—to childhood obesity through the use of BMI among 2-year-old Hispanic American children. I also evaluated clinical, environmental, and social risk factors for childhood obesity. The data for my study was obtained from the Inova Translational Medicine Institute. To my knowledge, this is the first study to examine the influence of ancestry on BMI among admixed Hispanic American children at age 2. Results of the analysis showed that EUR**,** AMR, and AFR were not associated with BMI at this age. Nevertheless, evaluating clinical and social risk factors showed that birth weight was

positively associated with BMI, and that higher maternal education levels represented a protective factor against higher BMI.

Interpretation of the Findings, Comparison, and Synthesis of Other Research Studies

The goal of this chapter is to discuss and synthesize the answers to the research questions and evaluate whether results confirm, disconfirm, or extend knowledge about the relationships between genetic admixture proportion, obesity, and BMI. The research questions asked whether associations existed between genetic certain backgrounds (EUR, AMR, or AFR) and BMI for 2-year-old Hispanic American children. Linear regression models were calculated to evaluate the relationship between BMI and genetic admixture proportions. No associations were found for BMI and EUR, BMI and AMR, or BMI and AFR. These results demonstrated that ancestral genetic background was not related to BMI for Hispanic American children at age 2. However, evaluating clinical, social, and environmental factors showed that birth weight was positively associated with BMI, and that a higher level of maternal education was a protective factor against higher BMI.

Many features of human anatomy and physiology are related to evolution; therefore, the principles of evolutionary biology may provide new insights into childhood obesity. Common polygenic obesity observed in adults originated during childhood or even in utero (Barker, 2012; Gluckman, Hanson, Cooper, & Thornburg, 2008). As such, it is important to understand how the association of genetic admixture and BMI that was discovered for adult populations operates in early childhood.

Obesity among children and adults has notably increased over recent decades and represents a major global health problem. Both genetic and environmental factors contribute to the complex etiology of polygenic obesity. Genome-wide association studies have identified about 100 loci associated with BMI; however, these variants only explain 2.7% of the variation in BMI. Therefore, many other variants remain unidentified due to small sample sizes; genetic heterogeneity; or epigenetic, gene-gene, or gene-environment interactions (Locke et al., 2015).

Previous studies reported that ancestral genetic background contributes to racial or ethnic differences in body composition (Cardel et al., 2011). In my analysis, EUR, AMR, and AFR admixture proportions were not associated with BMI among Hispanic American children at age 2. Individuals of mixed ancestral background have been found to have an above-average BMI, suggesting that differences in ancestral background may partially explain ethnic differences in the prevalence of obesity (Fernandez et al., 2013). Evaluating early childhood obesity among admixed Hispanic children provides an opportunity to understand the influence of ancestry genetic background on obesity while there has been less exposure to environmental factors. Herrera et al. (2011) argued that common obesity is heritable, but that it is very challenging to detect genetic variants that contribute to one's susceptibility to obesity. Furthermore, environmental exposures may affect epigenetic markers and result in obesity. Norden-Krichmar et al. (2014) explored the influence of the degree of AMR admixture on BMI in Native Americans. According to this study, genetic factors may explain some of the variations in obesity among Native Americans. Comuzzie et al. (2012) investigated the genetics of obesity among Hispanic

children and found novel genes with unknown function in obesity pathogenesis. However, their study has reported the variants identified are likely not the actual causal variants.

Contrary to my findings in 2-year-old children, in several studies conducted in the adult population, a positive association was observed between BMI and AFR, and similar to my study results, no association was observed between AMR and obesity (Cheng et al., 2010; Klimentidis et al., 2009). While previous studies have shown an association between AFR and obesity, my study finding can be somewhat limited by the small average proportion of AFR (0.08(SD ± 0.07)) in my study population. Furthermore, unlike to my study finding, in examining the relationship between EUR ancestries on obesity-related traits, EUR was found to have a protective effect (Cheng et al., 2010); this protective effect of EUR admixture was also demonstrated among Native-American college students (Klimentidis et al., 2009), suggesting the protective influence of EUR genetic background on low energy expenditure or more food consumption may vary with age, with less or no influence in very early childhood.

The development of obesity due to genes associated with hormones and neurotransmitters that regulate appetite and energy expenditure may happen later in life. The findings of this study emphasize the contribution of social and clinical factors to childhood BMI; specifically, maternal education levels and birth weight. This result is encouraging, as many of the social and environmental factors are modifiable, unlike an individual's ancestry background and genetic makeup.

Limitations of the Study

Data initially was collected for the "The First 1,000 Days of Life Study"; my study was a retrospective study using secondary/archival data from "The First 1,000 Days of Life Study." Some of the variables—including dietary values and anthropometrics—were parentally reported, therefore, differential misclassification likely biasing the results towards the null and information bias may have occurred during data collection, and internal validity may be jeopardized. Another threat to validity for this study was selection bias: firstly, parents who agreed to participate in the genomic study may have had different demographics than the ones who did not agree to participate in a longitudinal genomic study; secondly, parents who did not complete longitudinal surveys may have had a particular problem affecting this cross-sectional study. This is a particular problem when the characteristics of nonresponders differ from responders (Shepherd, 1998).

This study was limited to 2-year-old Hispanic American children residing in the Washington, DC, metropolitan area who have been recruited into "The First 1,000 Days of Life Study," and whose parents had completed longitudinal surveys. Although the children in this study were likely representative of Hispanic American ancestry admixture in the Washington DC metropolitan area, the result may not be generalizable to different acculturation and assimilation processes in other regions of the United States.

Recommendations

Childhood obesity is a serious problem in the United States, putting kids at risk for poor health. Despite recent declines in the prevalence of obesity in children aged 2-5

years, obesity amongst all children—particularly Hispanics and African Americans—is still too high. The prevalence of childhood obesity is higher than the Healthy People 2020 goal of 14.5%, and it is disproportionately higher for Hispanic children (Ogden, 2015). Hispanics are expected to represent 24% of U.S. population by 2050 (Lee, 2010). The Hispanic population is genetically diverse, representing a heterogeneous mix of EUR, AFR, and AMR ancestry (Gonzalez Burchard et al., 2005). The use of a single Hispanic or Latino ethnic category is insufficient for characterizing genetic background and disease prevalence (Lara et al., 2006). Additional longitudinal research with a larger study population should further explore the influence of genetic admixture on childhood obesity. Also, the development of educational policies and programs for Hispanic parents may help to decrease childhood obesity among Hispanic American children and eliminate disparity. Larger sample sizes and targeted primary data collection studies will allow for a more comprehensive investigation of disparities in childhood obesity rates, giving the opportunity for targeted and personalized anticipatory guidance to reduce obesity rates in teenage and adult populations.

Social Change Implication

Hispanic children have a higher prevalence of obesity compared to non-Hispanic White children in the United States (Ogden et al., 2014). However, knowledge about the association of obesity and children's ancestry is lacking, and this issue has not been well studied among Hispanic American children from different ancestry backgrounds. A deeper understanding of the associations between obesity and genetic admixture proportions of Hispanic American children—along with other social and clinical

factors—may help to better explain the prevalence and etiology of childhood obesity among admixed Hispanic American children.

There was not any association between children's BMI and admixture proportion of any ancestry background; however, this knowledge is beneficial because it differs from studies of adult populations that suggested a stronger relationship between genetics and obesity. The information gained from this study has the potential to create positive social change by developing preventions that targets modifiable childhood risk factors. Targeted and precise preventions can help to decrease childhood obesity, lower disparities in obesity rates for different groups, and lower the consequences of obesity over time among Hispanic American children as well as other so called "high risk" populations.

Conclusion

The association of genetic admixture and BMI may vary by age. Development of polygenic obesity involves genetic and nongenetic factors. Thus, the influence of hormones and neurotransmitters that regulate appetite and energy expenditure may change over time, and further studies are needed to investigate the role of ancestry genetic background on BMI by age to determine when and how this relationship develops. Understanding the genetic architecture can help prevention and treatment of obesity and will have fundamental implications for other diseases later in life. However, knowing that ancestry is not the main cause of disparities in rates of early childhood obesity is encouraging, as many of the social and environmental factors are modifiable. I cannot conclude that 2-year-old Hispanic American children's genetic admixture

proportion is associated with BMI; future research may begin to examine more detailed genetic markers in larger populations and understand the influence of hormones and neurotransmitters that regulate appetite and energy expenditure as children age. However, these results suggest an opportunity for targeted and personalized anticipatory guidance to reduce rates of childhood obesity for the Hispanic American population.

References

The 1,000 Genomes Project Consortium. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature, 491*(7422), 56–65. http://doi.org/10.1038/nature11632

The 1,000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68–74. http://doi.org/10.1038/nature15393

Arthur, R., Schulz-Trieglaff, O., Cox, A. J., & O'Connell, J. (2016). AKT: Ancestry and kinship toolkit. *Bioinformatics*, *33*(1), 142–144. doi:10.1093/bioinformatics/btw576

Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E., Kang, H. M., Korbel, J. O., . . . Albers, C. A. (2015). A global reference for human genetic variation. *Nature, 526*, 68–74.

Bacha, F., & Gidding, S. S. (2016). Cardiac abnormalities in youth with obesity and type 2 diabetes. *Current Diabetes Reports, 16*(7), 62.

Barker, D. J. (2012). Developmental origins of chronic disease. *Public Health, 126*(3), 185–189.

Bass, R., & Eneli, I. (2015). Severe childhood obesity: An under-recognised and growing health problem. *Postgraduate Medical Journal, 91*(1081), 639–645.

Batis, C., Hernandez-Barrera, L., Barquera, S., Rivera, J. A., & Popkin, B. M. (2011). Food acculturation drives dietary differences among Mexicans, Mexican Americans, and non-Hispanic Whites–3. *The Journal of Nutrition, 141*(10), 1898–1906.

Bryc, K., Durand, E. Y., Macpherson, J. M., Reich, D., & Mountain, J. L. (2015). The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *American Journal of Human Genetics, 96*(1), 37–53. http://doi.org/10.1016/j.ajhg.2014.11.010

Cardel, M., Higgins, P., Willig, A., Keita, A., Casazza, K., Gower, B., & Fernández, J. (2011). African genetic admixture is associated with body composition and fat distribution in a cross-sectional study of children. *International Journal of Obesity, 35*(1), 60–65. http://doi.org/10.1038/ijo.2010.203

Cecil, J. E., Tavendale, R., Watt, P., Hetherington, M. M., & Palmer, C. N. (2008). An obesity-associated FTO gene variant and increased energy intake in children. *New England Journal of Medicine, 359*(24), 2558–2566.

Centers for Disease Control and Prevention (CDC). (2001a). Data table of BMI-for-age charts (males). Retrieved from www.cdc.gov/growthcharts/html_charts /bmiagerev.htm#males

Centers for Disease Control and Prevention. (2001b). Data table of BMI-for-age charts (females). Retrieved from https://www.cdc.gov/growthcharts/html_charts /bmiagerev.htm#females

Centers for Disease Control and Prevention. (2012). Behavioral risk factor surveillance system questionnaire. Retrieved from https://www.cdc.gov/brfss/questionnaires /pdf-ques/2012_brfss.pdf

Centers for Disease Control and Prevention. (2014). Calculating BMI using the English system. Retrieved from https://www.cdc.gov/nccdphp/dnpao/growthcharts /training/bmiage/page5_2.html

Centers for Disease Control and Prevention. (2016a). Childhood obesity causes & consequences. Retrieved from https://www.cdc.gov/obesity /childhood/causes.html

Centers for Disease Control and Prevention. (2016b). Defining childhood obesity. Retrieved from https://www.cdc.gov/obesity/childhood/defining.html

Centers for Disease Control and Prevention. (2017). Childhood obesity facts. Retrieved from https://www.cdc.gov/obesity/data/childhood.html

Chen, W., Brehm, J. M., Boutaoui, N., Soto-Quiros, M., Avila, L., Celli, B. R., . . . Celedón, J. C. (2014). Native American ancestry, lung function, and COPD in Costa Ricans. *Chest, 145*(4), 704–710. http://doi.org/10.1378/chest.13-1308

Cheng, C. Y., Reich, D., Coresh, J., Boerwinkle, E., Patterson, N., Li, M., . . . Kao, W. H. (2010). Admixture mapping of obesity-related traits in African Americans: The Atherosclerosis Risk in Communities (ARIC) study. *Obesity, 18*(3), 563–572.

Cheung, W. W., & Mao, P. (2012). Recent advances in obesity: genetics and beyond. *International Scholarly Research Notices Endocrinology*, 2012. http://dx.doi.org/10.5402/2012/536905

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. NJ: Lawrence Erlbaum.

Colby, S. L., & Ortman, J. M. (2015). *Projections of the size and composition of the U.S. population: 2014 to 2060, Current Population Reports, P25-1143*. Washington, DC: U.S. Census Bureau. Retrieved from https://www.census.gov/content/dam /Census/library/publications/2015/demo/p25-1143.pdf

Comuzzie, A. G., Cole, S. A., Laston, S. L., Voruganti, V. S., Haack, K., Gibbs, R. A., & Butte, N. F. (2012). Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. *PloS One, 7*(12), e51954.

Cote, A. T., Harris, K. C., Panagiotopoulos, C., Sandor, G. G., & Devlin, A. M. (2013). Childhood obesity and cardiovascular dysfunction. *Journal of the American College of Cardiology, 62*(15), 1309–1319.

Deng, L., Ruiz-Linares, A., Xu, S., & Wang, S. (2016). Ancestry variation and footprints of natural selection along the genome in Latin American populations. *Scientific Reports, 6*, 21766. doi:10.1038/srep21766

Divers, J., Palmer, N. D., Lu, L., Register, T. C., Carr, J. J., Hicks, P. J., . . . Freedman, B. I. (2013). Admixture mapping of coronary artery calcified plaque in African Americans with type 2 diabetes mellitus. Circulation. *Cardiovascular Genetics, 6*(1), 97–105. http://doi.org/10.1161/CIRCGENETICS.112.964114

Dobbs, R., Sawers, C., Thompson, F., Manyika, J., Woetzel, J. R., Child, P., . . . Spatharou, A. (2014). Overcoming obesity: An initial economic analysis. Jakarta, Indonesia: McKinsey Global Institute.

Dubois, L., Kyvik, K. O., Girard, M., Tatone-Tokuda, F., Pérusse, D., Hjelmborg, J., . . . Martin, N. G. (2012). Genetic and environmental contributions to weight, height, and BMI from birth to 19 years of age: An international study of over 12,000 twin pairs. *PLOS one, 7*(2), e30153.

Ellaway, A., Anderson, A., & Macintyre, S. (1997). Does area of residence affect body size and shape? *International Journal of Obesity, 21*(4), 304.

Ernst, A. F., & Albers, C. J. (2017). Regression assumptions in clinical psychology research practice—a systematic review of common misconceptions. *PeerJ*, *5*, e3323. http://doi.org/10.7717/peerj.3323

Faith, M. S., & Kral, T. V. (2006). Social environmental and genetic influences on obesity and obesity-promoting behaviors: Fostering research integration. In Hernandez, L. M., & Blazer, D. G. (Eds.,), *Genes, behavior, and the social environment: Moving beyond the nature/nurture debate*. Washington, DC: National Academies Press.

Fall, T., & Ingelsson, E. (2014). Genome-wide association studies of obesity and metabolic syndrome. *Molecular and Cellular Endocrinology, 382*(1), 740–757.

Farooqi, I. S., & O'Rahilly, S. (2006). Genetics of obesity in humans. *Endocrine Reviews, 27*(7), 710–718.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). *G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Fernandez, J. R., Pearson, K. E., Kell, K. P., & Brown, M. M. B. (2013). Genetic admixture and obesity: Recent perspectives and future applications. *Human Heredity, 75*(0). http://doi.org/10.1159/000353180

Fiorito, L. M., Marini, M., Francis, L. A., Smiciklas-Wright, H., & Birch, L. L. (2009). Beverage intake of girls at age 5 y predicts adiposity and weight status in childhood and adolescence. *The American Journal of Clinical Nutrition, 90*(4), 935–942. http://doi.org/10.3945/ajcn.2009.27623

Freedman, D. S., Horlick, M., & Berenson, G. S. (2013). A comparison of the Slaughter skinfold-thickness equations and BMI in predicting body fatness and cardiovascular disease risk factor levels in children. *The American Journal of Clinical Nutrition, 98*(6), 1417–1424.

Fryar, C. D., Carroll, M. D., & Ogden, C. L. (2012). *Prevalence of overweight, obesity, and extreme obesity among adults: United States, trends 1960–1962 through 2009–2010.* Retrieved from https://www.cdc.gov/nchs/data/hestat /obesity_adult_11_12/obesity_adult_11_12.htm

Galanter, J. M., Gignoux, C. R., Oh, S. S., Torgerson, D., Pino-Yanes, M., Thakur, N., . . . Zaitlen, N. (2017). Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *eLife, 6*, e20532. http://doi.org/10.7554/eLife.20532

Garrow, J. S., & Webster, J. (1985). Quetelet's index (W/H2) as a measure of fatness. *International Journal of Obesity, 9*(2), 147–153.

Garver, W. S., Newman, S. B., Gonzales-Pacheco, D. M., Castillo, J. J., Jelinek, D., Heidenreich, R. A., & Orlando, R. A. (2013). The genetics of childhood obesity and interaction with dietary macronutrients. *Genes & Nutrition, 8*(3), 271–287. http://doi.org/10.1007/s12263-013-0339-5

Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of in utero and early-life conditions on adult health and disease. *New England Journal of Medicine, 359*(1), 61–73.

Gluckman, P. D., Low, F. M., Buklijas, T., Hanson, M. A., & Beedle, A. S. (2011). How evolutionary principles improve the understanding of human health and disease. *Evolutionary Applications, 4*(2), 249–263. http://doi.org/10.1111/j.1752-4571.2010.00164.x

Gonzalez Burchard, E., Borrell, L. N., Choudhry, S., Naqvi, M., Tsai, H.-J., Rodriguez-Santana, J. R., . . . Risch, N. (2005). Latino populations: A unique opportunity for the study of race, genetics, and social environment in epidemiological research. *American Journal of Public Health, 95*(12), 2161–2168. http://doi.org/10.2105/AJPH.2005.068668

Grant, C., & Osanloo, A. (2014). Understanding, selecting, and integrating a theoretical framework in dissertation research: Creating the blueprint for your "house." *Administrative Issues Journal, 4*(2), 4.

Griffiths, L. J., Hawkins, S. S., Cole, T. J., & Dezateux, C. (2010). Risk factors for rapid weight gain in preschool children: Findings from a UK-wide prospective study. *International Journal of Obesity, 34*(4), 624.

Grundy, S. M. (1998). Multifactorial causation of obesity: Implications for prevention. *The American Journal of Clinical Nutrition, 67*(3), 563S–572S.

Hall, B. K. (2003). Evo-devo: Evolutionary developmental mechanisms. *International Journal of Developmental Biology, 47*(7-8), 491–495.

Hall, B. K. (2012). Evolutionary developmental biology (Evo-Devo): Past, present, and future. *Evolution: Education and outreach, 5*(2), 184–193.

Hammond, R. A., & Levine, R. (2010). The economic impact of obesity in the United States. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy, 3*, 285–295. http://doi.org/10.2147/DMSOTT.S7384

Hassanein, M. T., Lyon, H. N., Nguyen, T. T., Akylbekova, E. L., Waters, K., Lettre, G., . . . Butler, J. L. (2010). Fine mapping of the association with obesity at the FTO locus in African-derived populations. *Human Molecular Genetics, 19*(14), 2907–2916.

Hazrati, S., Hourigan, S. K., Waller, A., Yui, Y., Gilchrist, N., Huddleston, K., & Niederhuber, J. (2016). Investigating the accuracy of parentally reported weights and lengths at 12 months of age as compared to measured weights and lengths in a longitudinal childhood genome study. *BMJ Open, 6*(8), e011653.

He, M., Harris, S., Piché, L., & Beynon, C. (2009). Understanding screen-related sedentary behavior and its contributing factors among school-aged children: A social-ecologic exploration. *American Journal of Health Promotion, 23*(5), 299–308.

Hernandez-Suarez, G., Sanabria, M. C., Serrano, M., Herran, O. F., Perez, J., Plata, J. L., . . . Tenesa, A. (2014). Genetic ancestry is associated with colorectal adenomas and adenocarcinomas in Latino populations. *European Journal of Human Genetics, 22*(10), 1208–1216. doi:10.1038/ejhg.2013.310

Herrera, B. M., Keildson, S., & Lindgren, C. M. (2011). Genetics and epigenetics of obesity. *Maturitas, 69*(1), 41–49. http://doi.org/10.1016/j.maturitas.2011.02.018

Hesketh, K., Waters, E., Green, J., Salmon, L., & Williams, J. (2005). Healthy eating, activity and obesity prevention: A qualitative study of parent and child perceptions in Australia. *Health Promotion International, 20*(1), 19–26.

Higgins, P. B., Fernández, J. R., Goran, M. I., & Gower, B. A. (2005). Early ethnic difference in insulin-like growth factor-1 is associated with African genetic admixture. *Pediatric Research, 58*(5), 850–854.

Hill, K. (1993). Life history theory and evolutionary anthropology. *Evolutionary Anthropology: Issues, News, and Reviews, 2*(3), 78–88.

Hoggart, C. J., Parra, E. J., Shriver, M. D., Bonilla, C., Kittles, R. A., Clayton, D. G., & McKeigue, P. M. (2003). Control of confounding of genetic associations in stratified populations. *American Journal of Human Genetics, 72*(6), 1492–1504.

Hu, F. (2008). Genetic predictors of obesity. *Obes Epidemiol,* 437–460. 10.1093/acprof:oso/9780195312911.001.0001

Hunninghake, G. M., Weiss, S. T., & Celedón, J. C. (2006). Asthma in Hispanics. *American Journal of Respiratory and Critical Care Medicine, 173*(2), 143–163. http://doi.org/10.1164/rccm.200508-1232SO

Isasi, C. R., Rastogi, D., & Molina, K. (2016). Health issues in Hispanic/Latino youth. *Journal of Latina/o Psychology, 4*(2), 67.

Kahn, H. S., Tatham, L. M., Pamuk, E. R., & Heath, C. W., Jr. (1998). Are geographic regions with high income inequality associated with risk of abdominal weight gain? *Social Science & Medicine, 47*(1), 1–6.

Kappil, M., Wright, R. O., & Sanders, A. P. (2016). Developmental origins of common disease: Epigenetic contributions to obesity. *Annual Review of Genomics and Human Genetics, 17,* 177–192. http://doi.org/10.1146/annurev-genom-090314-050057

Kim, J., & Seo, B. S. (2013). How to calculate sample size and why. *Clinics in Orthopedic Surgery, 5*(3), 235–242. http://doi.org/10.4055/cios.2013.5.3.235

Klimentidis, Y. C., Miller, G. F., & Shriver, M. D. (2009). The relationship between European genetic admixture and body composition among Hispanics and Native Americans. *American Journal of Human Biology, 21*(3), 377–382.

Kühnen, P., Handke, D., Waterland, R. A., Hennig, B. J., Silver, M., Fulford, A. J., . . . Hinney, A. (2016). Interindividual variation in DNA methylation at a putative POMC metastable epiallele is associated with obesity. *Cell metabolism, 24*(3), 502–509.

Lara, M., Akinbami, L., Flores, G., & Morgenstern, H. (2006). Heterogeneity of childhood asthma among Hispanic children: Puerto Rican children bear a disproportionate burden. *Pediatrics, 117*(1), 43–53.

Lin, S. S., & Kelsey, J. L. (2000). Use of race and ethnicity in epidemiologic research: Concepts, methodological issues, and suggestions for research. *Epidemiologic Reviews, 22*(2), 187–202.

Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., . . . Croteau-Chonka, D. C. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature, 518*(7538), 197.

Loos, R. J., Lindgren, C. M., Li, S., Wheeler, E., Zhao, J. H., Prokopenko, I., . . . Mohike, K. L. (2008). Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nature Genetics, 40*(6), 768.

Martínez, J. A., Milagro, F. I., Claycombe, K. J., & Schalinske, K. L. (2014). Epigenetics in adipose tissue, obesity, weight loss, and diabetes. *Advances in Nutrition: An International Review Journal, 5*(1), 71–81.

Maskarinec, G., Morimoto, Y., Jacobs, S., Grandinetti, A., Mau, M. K., & Kolonel, L. N. (2016). Ethnic admixture affects diabetes risk in Native Hawaiians: The multiethnic cohort. *European Journal of Clinical Nutrition, 70*(9), 1022–1027. http://doi.org/10.1038/ejcn.2016.32

McCrum-Gardner, E. (2010). Sample size and power calculations made simple. *International Journal of Therapy and Rehabilitation, 17*(1), 10–14.

McGavock, J. M., Torrance, B. D., McGuire, K. A., Wozny, P. D., & Lewanczuk, R. Z. (2009). Cardiorespiratory fitness and the risk of overweight in youth: The Healthy Hearts Longitudinal Study of Cardiometabolic Health. *Obesity, 17*(9), 1802–1807.

McKeigue, P. M. (2005). Prospects for admixture mapping of complex traits. *The American Journal of Human Genetics, 76*(1), 1–7.

Mersha, T. B. (2015). Mapping asthma-associated variants in admixed populations. *Frontiers in Genetics, 6,* 292. doi:10.3389/fgene.2015.00292

Mersha, T. B., & Abebe, T. (2015). Self-reported race/ethnicity in the age of genomic research: Its potential impact on understanding health disparities. *Human Genomics, 9*(1), 1. http://doi.org/10.1186/s40246-014-0023-x

Must, A., Phillips, S. M., Naumova, E. N. (2012). Occurrence and timing of childhood overweight and mortality: Findings from the Third Harvard Growth Study. *The Journal of Pediatrics, 160*(5), 743–750.

National Heart Lung and Blood Institute. (2016). Aim for a healthy weight. Available from http://www.nhlbi.nih.gov/health/educational/lose_wt/

Norden-Krichmar, T. M., Gizer, I. R., Libiger, O., Wilhelmsen, K. C., Ehlers, C. L., & Schork, N. J. (2014). Correlation analysis of genetic admixture and social identification with body mass index in a Native American community. *American Journal of Human Biology: The Official Journal of the Human Biology Council, 26*(3), 347–360. doi:10.1002/ajhb.22521

Nuttall, F. Q. (2015). Body mass index: Obesity, BMI, and health: A critical review. *Nutrition Today, 50*(3), 117.

Ogden, C. L., Carroll, M. D., Fryar, C. D., & Flegal, K. M. (2015). *Prevalence of obesity among adults and youth: United States, 2011–2014*. Washington, DC: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2014). Prevalence of childhood and adult obesity in the United States, 2011–2012. *JAMA, 311*(8), 806–814.

Ogden, C. L., & Flegal, K. M. (2010). Changes in terminology for childhood overweight and obesity. *Age, 12*(12).

Ordoñez, G., Romero, S., Orozco, L., Pineda, B., Jimenez-Morales, S., Nieto, A., . . . Sotelo, J. (2015). Genomewide admixture study in Mexican Mestizos with multiple sclerosis. *Clinical Neurology and Neurosurgery, 130,* 55–60. doi:10.1016/j.clineuro.2014.11.026

Pan, L. (2016). Trends in obesity among participants aged 2–4 years in the Special Supplemental Nutrition Program for Women, Infants, and Children—United States, 2000–2014. *Morbidity and Mortality Weekly Report, 65*.

Peterson, T., & Müller, G. B. (2016). Phenotypic novelty in EvoDevo: The distinction between continuous and discontinuous variation and its importance in evolutionary theory. *Evolutionary Biology, 43,* 314–335. http://doi.org/10.1007/s11692-016-9372-9

Pigeyre, M., Yazdi, F. T., Kaur, Y., & Meyre, D. (2016). Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity. *Clinical Science, 130*(12), 943–986.

Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J. S., Albrechtsen, A., Moltke, I., . . . Bertalan, M. (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature, 463*(7282), 757–762.

Ricks-Santi, L. J., Apprey, V., Mason, T., Wilson, B., Abbas, M., Hernandez, W., . . . Ahaghotu, C. (2012). Identification of genetic risk associated with prostate cancer using ancestry informative markers. *Prostate Cancer and Prostatic Diseases, 15*(4), 359–364. http://doi.org/10.1038 /pcan.2012.19

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., & Feldman, M. W. (2002). Genetic structure of human populations. *Science, 298*(5602), 2381–2385.

Sahoo, K., Sahoo, B., Choudhury, A. K., Sofi, N. Y., Kumar, R., & Bhadoria, & A. S. (2015). Childhood obesity: Causes and consequences. *Journal of Family Medicine and Primary Care, 4*(2), 187–192. http://doi.org/10.4103/2249-4863.154628

Salzano, F. M., & Sans, M. (2014). Interethnic admixture and the evolution of Latin American populations. *Genetics and Molecular Biology, 37*(1), 151–170.

Scholtens, S., Brunekreef, B., Visscher, T. L., Smit, H. A., Kerkhof, M., Jongste, J. C. D., . . . Wijga, A. H. (2006). Reported versus measured body weight and height of 4-year-old children and the prevalence of overweight. *The European Journal of Public Health, 17*(4), 369–374. https://doi.org/10.1093.eurpub/ckl253

Shaffer, J. R., Kammerer, C. M., Reich, D., McDonald, G., Patterson, N., Goodpaster, B., . . . Health ABC Study (2007). Genetic markers for ancestry are correlated with body composition traits in older African Americans. *Osteoporosis International, 18*(6), 733–741.

Shepherd, C. K., Power, K. G., & Carter, H. (1998). Characteristics of responders and non-responders in an infant feeding study. *Journal of Public Health, 20*(3), 275–280.

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics, 195*(3), 693–702.

Speakman, J. R. (2013). Evolutionary perspectives on the obesity epidemic: Adaptive, maladaptive, and neutral viewpoints. *Annual Review of Nutrition, 33,* 289–317.

Statistics Solutions. (2013). *Data analysis plan: Multiple linear regression.* Retrieved from http://www.statisticssolutions.com/academic-solutions/member-resources/member-profile/data-analysis-plan-templates/data-analysis-plan-multiple-linear-regression/

Stryjecki, C., Alyass, A., & Meyre, D. (2017). Ethnic and population differences in the genetic predisposition to human obesity. *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity, 19*(1), 62–80. doi:10.1111/obr.12604

Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: Provocative insights from epigenomics. *Nature Reviews Genetics, 9*(6), 465.

Taveras, E. M., Gillman, M. W., Kleinman, K. P., Rich-Edwards, J. W., & Rifas-Shiman, S. L. (2013). Reducing racial/ethnic disparities in childhood obesity: The role of early life risk factors. *JAMA Pediatrics, 167*(8). 10.1001/jamapediatrics.2013.85.

Taveras, E. M., Rifas-Shiman, S. L., Oken, E., Gunderson, E. P., & Gillman, M. W. (2008). Short sleep duration in infancy and risk of childhood overweight. *Archives of Pediatrics & Adolescent Medicine, 162*(4), 305–311.

The First 1,000 Days of Life and Beyond. (2018). Retrieved from https://www.inova.org/itmi/our-research-studies/first-1000-days-of-life

Tofthagen, C. (2012). Threats to validity in retrospective studies. *Journal of the Advanced Practitioner in Oncology, 3*(3), 181–183.

Tremmel, M., Gerdtham, U.-G., Nilsson, P. M., & Saha, S. (2017). Economic burden of obesity: A systematic literature review. *International Journal of Environmental Research and Public Health, 14*(4), 435. http://doi.org/10.3390/ijerph14040435

Williams, D. E., Cadwell, B. L., Cheng, Y. J., Cowie, C. C., Gregg, E. W., Geiss, L. S., . . . Imperatore, G. (2005). Prevalence of impaired fasting glucose and its relationship with cardiovascular disease risk factors in U.S. adolescents, 1999–2000. *Pediatrics, 116*(5), 1122–1126.

Yudell, M., Roberts, D., DeSalle, R., & Tishkoff, S. (2016). Taking race out of human genetics. *Science, 351*(6273), 564–565.

Zint, M. (2018). *Power analysis, statistical significance, & effect size.* Retrieved from http://meera.snre.umich.edu/power-analysis-statistical-significance-effect-size

Ziv, E., John, E. M., Choudhry, S., Kho, J., Lorizio, W., Perez-Stable, E. J., & Burchard, E. G. (2006). Genetic ancestry and risk factors for breast cancer among Latinas in the San Francisco Bay Area. *Cancer Epidemiology, Biomarkers & Prevention15*(10), 1878–1885. http://doi.org/10.1158/1055-9965.EPI-06-0092