UNIVERSITY OF CALIFORNIA

Los Angeles

Evaluating Community-level Initiatives to Address Early Childhood Obesity in Los Angeles County:

An Innovative Application of Machine Learning Methods to Community Health Research

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Community Health Sciences

by

Shelley Jung

2020

ABSTRACT OF THE DISSERTATION

Evaluating Community-level Initiatives to

Address Early Childhood Obesity in Los Angeles County:

An Innovative Application of Machine Learning Methods

to Community Health Research

by

Shelley Jung

Doctor of Philosophy in Community Health Sciences

University of California, Los Angeles, 2020

Professor May-Choo Wang, Chair

Obesity prevalence among children in the United States has almost doubled in the past three decades. To address the rising rates of obesity, community-level interventions have been implemented. However, it remains unclear whether resources are reaching all communities with need or what factors determine the allocation of scarce resources. Communities most burdened by obesity should be prioritized for intervention. However, due to a lag time in data availability, current obesity estimates needed to identify communities with the greatest needs for intervention are not available. Furthermore, evidence demonstrating the contribution of place-based interventions to changes in population-level rates of childhood obesity has been limited.

A database of interventions tackling obesity in Los Angeles County since 2003 was created. Neighborhood-level intervention data was linked with neighborhood-level obesity prevalence and sociodemographic data. Generalized linear models with a Gamma distribution and log link were run to examine the allocation of resources for obesity prevention across communities based on their obesity prevalence and sociodemographic characteristics. Machine learning algorithms were used to build models predicting future prevalence of neighborhood-level obesity rates using existing neighborhood sociodemographic and obesity data. Machine learning algorithms were also applied to build a model to estimate neighborhood-level prevalence of obesity under no intervention, which was used to create a counterfactual comparison group. This model was applied to neighborhoods that received intervention(s) in a given year to estimate what their obesity prevalence would have been under no intervention. We ran fixed-effects models to examine the relationship between various types of obesity-related interventions and change in obesity prevalence.

Neighborhoods with more poverty and a higher proportion of Black or Hispanic residents were more likely to receive obesity-related interventions. We also demonstrated that future prevalence of neighborhood-level obesity could be reasonably predicted using the most recent sociodemographic and obesity data available. Finally, we found that neighborhoods that received more obesity-related interventions saw greater declines in obesity prevalence. In particular, neighborhoods that received multicomponent interventions were likelier to see greater declines in obesity prevalence. Macro-level interventions were more effective at reducing neighborhood-level prevalence of obesity than micro-level interventions.

The dissertation of Shelley Jung is approved.

Michael L. Prelip

Wendelin M. Slusser

Onyebuchi Aniweta Arah

Paul M. Ong

May-Choo Wang, Committee Chair

University of California, Los Angeles

2020

DEDICATION PAGE

To my loving parents and my supportive husband Daniel.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

ACKNOWLEDGEMENTS

I would like to thank May Wang for her continuous support over the past 5+ years. May, you have been more than a mentor and advisor to me— you are my public health mom. I am incredibly grateful for your kindness, wisdom, and support, and the generosity you have shown me since I met you in Singapore. Thank you for allowing me to explore my research interests. I could not have done this without your support.

I would also like to thank my committee members— Dr. Michael Prelip, Dr. Onyebuchi Arah, Dr. Wendelin Slusser, Dr. Paul Ong, and Dr. Leobardo Estrada (may you rest in peace). I am grateful to have had all of you serve on my committee. I have learned so much from each one of you, and you have all shaped me to be a better multidisciplinary researcher. All of your comments, differing perspectives, and guidance has really helped me to improve my research and frame my work in a way that can make an impact on public health policy.

To my friends at the Fielding School of Public Health, this journey would not have been the same without all of you. Thank you to Amy Chai, Adva Gadoth, Jee Won Park, Hayley Ashbaugh, Mahmoud Gaddoury, and Vahe Khachadourian for letting me join your study group— though I was an outsider, you welcomed me with open arms and I am grateful for that. I really would not have survived EPI M204 without meeting you all. My interest in causal inference was supported by Solomon Makgoeng and Roch Nianogo. Thank you both for being the best TAs—answering all of my many questions, explaining difficult concepts to me, and helping me to get a better handle on causal inference methods. Adva Gadoth, Anna Hing, and Joanna Calderon— our sushi nights, movie nights, and nights decorating cookies/cupcakes have kept me sane. I am grateful for all of those memories and cherish the times we spent together.

I would like to express my gratitude to my family— my parents Kyung Ok Min and Jae Kyu Jung, and my brother Jason Jung. Thank you for your prayers, your support, and your encouragement. 엄마 아빠, thank you for being so tough on me while I was growing up. I really would not be the person I am today without you, and I am incredibly grateful to have had you as parents. 엄마 아빠 사랑해요.

Finally, to my husband Daniel Yoo. You are truly my life partner, and I am grateful for the patience you have shown as I frantically worked towards finishing my dissertation. I really could not have done this without you— our evenings discussing statistics, your insightful feedback on my research, cooking dinner for us. You are my biggest cheerleader and I am so grateful for your encouragement throughout this entire process. I love you so much.

VITA

Shelley Jung

EDUCATION

2013	Master of Science in Public Health (MScPH) in Epidemiology McGill University, Montreal, QC
2011	Honors Bachelor of Science (HBSc) in Human Biology University of Toronto, Toronto, ON

PROFESSIONAL EXPERIENCE

2018-2019	Teaching Associate for CHS 212 and CHS 100 Department of Community Health Sciences, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, CA
2017-2019	Graduate Student Fellow Public Health Foundation Enterprises WIC Program, Irwindale, CA
2017-2018	Teaching Assistant for CHS 210 and CHS 100 Department of Community Health Sciences, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, CA
2015-2018	Graduate Student Researcher Department of Community Health Sciences, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, CA
2016-2017	Graduate Student Researcher Semel Healthy Campus Initiative Center, University of California, Los Angeles, Los Angeles, CA
2016	Data Mining Summer Fellow Public Health Foundation Enterprises WIC Program, Irwindale, CA
2016	Public Health Scholar Los Angeles County Department of Public Health, Los Angeles, CA
2013-2015	Research Associate College of Alice & Peter Tan, National University of Singapore, Singapore
2013	Consultant The Research Institute of the McGill University Health Centre, Kangiqsualujuuaq, QC

HONORS AND AWARDS

2019-2020	Dissertation Year Fellowship
2017-2018	Department Block Grant Opportunity Award
2017	Graduate Summer Research Mentorship
2016-2017	Juneal Marie Smith Fellowship in International Nutrition
2015-2017	Juneal Marie Smith Fellowship in International Nutrition

PEER-REVIEWED PUBLICATIONS

1. Jung S, Nobari TZ, Whaley SE. Breastfeeding outcomes among WIC-participating infants and their relationships to Baby-Friendly Hospital practices. Breastfeed Med. 2019 Jul/Aug;14(6):424-431. doi: 10.1089/bfm.2019.0004.

2. Khan FA, Fox GJ, Lee RS, Riva M, Benedetti A, Proulx JF, Jung S, Hornby K, Behr MA, Menzies D. Housing and tuberculosis in an Inuit village in northern Quebec: A case-control study. CMAJ Open. 2016 Sept 19;4(3):E496-E506. doi:10.9778/cmajo.20160049.

3. Fox GJ, Lee RS, Lucas M, Khan FA, Proulx JF, Hornby K, Jung S, Benedetti A, Behr MA, Menzies D. Inadequate diet is associated with acquiring Mycobacterium tuberculosis infection in an Inuit community. A case-control study. Ann Am Thorac Soc. 2015 Aug;12(8):1153-62. doi: 10.1513/AnnalsATS.201503-156OC.

SELECTED PRESENTATIONS

1. Jung S, Jiang L, Kamareddine L, Nobari TZ, Roper-Fingerhut H, Nianogo R, Prelip M, Wang J, Whaley SE, Wang MC. Analysis of Community-level Interventions Addressing Childhood Obesity in Los Angeles County. American Public Health Association (APHA) 2019 Annual Meeting & Expo, Philadelphia, PA, USA, November 2019. [Oral presentation: Community-Based Interventions Aimed at Reducing Childhood Obesity]

2. Jung S. Using Sociodemographic Characteristics to Predict Early Childhood Obesity Among Low-Income Children in Los Angeles County: An Application of Machine Learning Techniques. 2018 Society for Epidemiologic Research Annual Meeting, Baltimore, MD, USA, June 2018. [Poster Session 1]

3. Jung S. Mediation Analysis of Gestational Weight Gain and Diabetes in Relation to Childhood Obesity Risk. 2017 Society for Epidemiologic Research Annual Meeting, Seattle, WA, USA, June 2017. [Poster Session 3]

CHAPTER 1: Introduction

In the United States (US), childhood obesity is a major public health problem. Obesity prevalence among children has almost doubled in the past three decades,[1] and today, the country has some of the highest obesity rates in the world.[2] In the US, 18.5% of children aged 2 to 19 years, and 13.9% of children aged 2 to 5 years are obese.[3] Children who are obese are more likely to have poor mental and physical health, and are at increased risk of developing non-communicable diseases like diabetes and cardiovascular disease at an earlier age.[4-6] Reducing obesity risk, specifically during childhood, is a public health priority as children who are obese are more likely to remain obese into adulthood.[7,8]

Biologically, obesity is due to energy intake that exceeds energy expenditure, which is largely determined by diet and levels of physical activity.[9] The regulation of body weight, however, is a complex process that involves genetic, endocrine, behavioral, psychosocial, and environmental factors.[10] Addressing obesity has been very challenging, and interventions to address obesity that have focused solely on individuals have demonstrated limited success.[11-13] Consequently, approaches to health promotion have shifted towards a multifactorial approach guided by a socioecological framework[14] to address the myriad of factors that increase obesity risk.[15-17]

Over the last decade, more than $2 billion was pledged by the Centers for Disease Control and Prevention, the Robert Wood Johnson Foundation, and The California Endowment, alone, to target and help reverse the obesity epidemic through community programs, policies, and interventions.[18-20] Obesity is being addressed through place-based health promotion initiatives that are multilevel, multicomponent, and implemented through multiple sectors and settings of an entire community.[21-23] These place-based "whole of neighborhood" interventions seek to reduce obesity prevalence for entire populations. Many efforts have been made to evaluate such place-based interventions,[21-25] but their evaluation has been analytically challenging. Studies have shown that place-based interventions show some promise for preventing obesity, leading to modest reductions in population weight gain,[21,22,24] albeit the evidence has been mixed.[23,25] Based on the literature, the contribution of place-based interventions to changes in population-level rates of childhood obesity is limited.[26]

This dissertation attempts to contribute to the growing body of literature on place-based interventions for obesity control by:

(1) Examining the distribution of community-level interventions addressing early childhood obesity in Los Angeles County and understanding how resources for obesity prevention programs are allocated across communities;

(2) Determining whether existing sociodemographic and obesity prevalence estimates can be used to identify communities most burdened by obesity in a timely manner; and

(3) Examining the contribution of place-based interventions to declines in neighborhood-level rates of early childhood obesity, and identifying the types of interventions that produce the greatest reductions in neighborhood-level early childhood obesity.

In Chapter 2, we discuss the importance of early childhood obesity, and introduce the conceptual framework for investigating the various risk factors that contribute to the development of early childhood obesity. We also provide an overview of how the complexities of early childhood obesity have been addressed through public health interventions over the past few decades.

In Chapter 3, we describe the three data sources used in this dissertation.

In Chapter 4, we examine the distribution of community-level interventions addressing obesity in Los Angeles County and examine how resources for obesity prevention programs are allocated across communities.

In Chapter 5, we explore whether existing sociodemographic and obesity prevalence estimates can be used to identify, in a timely manner, communities most burdened by obesity. We demonstrate the strength of using existing data, specifically, the latest obesity prevalence estimates and sociodemographic data available, to predict community-level obesity burden, with the application of sophisticated machine learning modelling techniques.

In Chapter 6, we examine the contribution of place-based interventions to declines in neighborhood-level rates of early childhood obesity, and identify the types of interventions that produced the greatest reductions in neighborhood-level prevalence of early childhood obesity.

CHAPTER 2: Background and significance

Childhood obesity is one of the most serious public health challenges of the 21st century.[4] Obesity in childhood starts very early in life. Children who are obese are more likely to remain obese into adulthood,[7,8] and are more likely to develop non-communicable diseases like diabetes and cardiovascular diseases at an earlier age.[4] Obesity can affect a child's physical health through nearly every organ system including cardiovascular, endocrine, and pulmonary systems, and can cause functional limitations.[27-30] Certain comorbidities such as type 2 diabetes mellitus, which used to present in adulthood, are now seen in children with obesity.[30] Obesity can also affect children's psychosocial health through the development of poor self-esteem, eating disorders, and depression,[5,6] and has been found to affect educational attainment.[31]

Globally, it is estimated that 38 million infants and young children under the age of 5 were overweight or obese in 2019.[32] The United States (US) has the highest obesity rates among members countries of the Organisation for Economic Co-operation and Development (OECD),[2] with one in seven US children aged 2 to 5 years who were obese in 2015-2016.[3] Obesity is widespread, serious, and costly. Compared to normal weight children who maintain a normal weight through adulthood, the lifetime direct medical cost of childhood obesity is $19,000.[33] The rising costs of health care in the US cannot be addressed without addressing obesity.

Definition and operationalization of obesity

Obesity refers to a state of excessive accumulation of adipose tissue or body fat.[34] Since directly measuring adiposity or body fat poses a number of practical challenges in the field, researchers often rely on proxies derived from simple, inexpensive anthropometric measures, such as height, weight, and waist circumference.[35] Adiposity is usually measured by Body Mass Index (BMI) since BMI correlates highly with adiposity in both children over 2 years of age and adults.[36] BMI is calculated by dividing weight in kilograms by height in meters squared (kg/m^2).[32] For adults, obesity is operationally defined by a BMI of 30 kg/m^2 or greater,[32] which correlates with increased risk of morbidity and mortality.[37-40] Among children, obesity in less well defined. Since children are continuously growing, anthropometric measures to determine obesity in children are confounded by natural age-related physiological variations in body proportion.[41,42] The World Health Organization (WHO) defines early childhood obesity (from birth to 5 years) as having a BMI > 3 standard deviations above the WHO growth standard median for obesity.[43] The International Obesity Task Force provides international BMI cut points by age and sex to define obesity in children aged 2 to 18.[44] The most common definition used by studies conducted in the US are based on the Centers for Disease Control and Prevention's (CDC) 2000 Growth Charts, developed from five national cross-sectional and longitudinal data sets in the US.[45] While there is no current consensus on the appropriate definition and measurement of obesity in very young children, obesity in children 2 to 19 years is generally defined as having a BMI that is ≥95th percentile of CDC's sex- and age-specific Growth Charts.[46] Adiposity in children is measured by BMI or weight-for-height (WH) expressed in terms of z-scores derived from growth reference values.[46] For this dissertation, we define obesity for young children as having a BMI ≥95th percentile of CDC's sex- and age-specific growth reference values.[46]

Early childhood obesity: A significant public health issue

Early childhood, or the preschool-aged years, is a critical developmental period early in the life course to focus obesity prevention efforts.[47] We refer to preschool-aged children as children between the ages of 2.0 and 5.0 (through the child's fifth birthday). Once childhood obesity is established, it tracks into adolescence, and persists into adulthood,[7,8] and is difficult to reverse through interventions.[11-13] Over half of obese children remain obese as adolescents, and around 80% of obese adolescents remain obese in adulthood.[8] Obese children and adolescents are about five times more likely to be obese in adulthood compared to children and adolescents who are not obese.[8]

Biologically, obesity is due to energy imbalance, that is, energy intake that exceeds energy expenditure. Energy intake is determined by diet, while energy expenditure is largely determined by levels of physical activity (PA).[9] Prevention efforts that focus on the preschool-aged years, when habits and behaviors related to eating and PA begin to develop[48] have the potential for lasting effects.

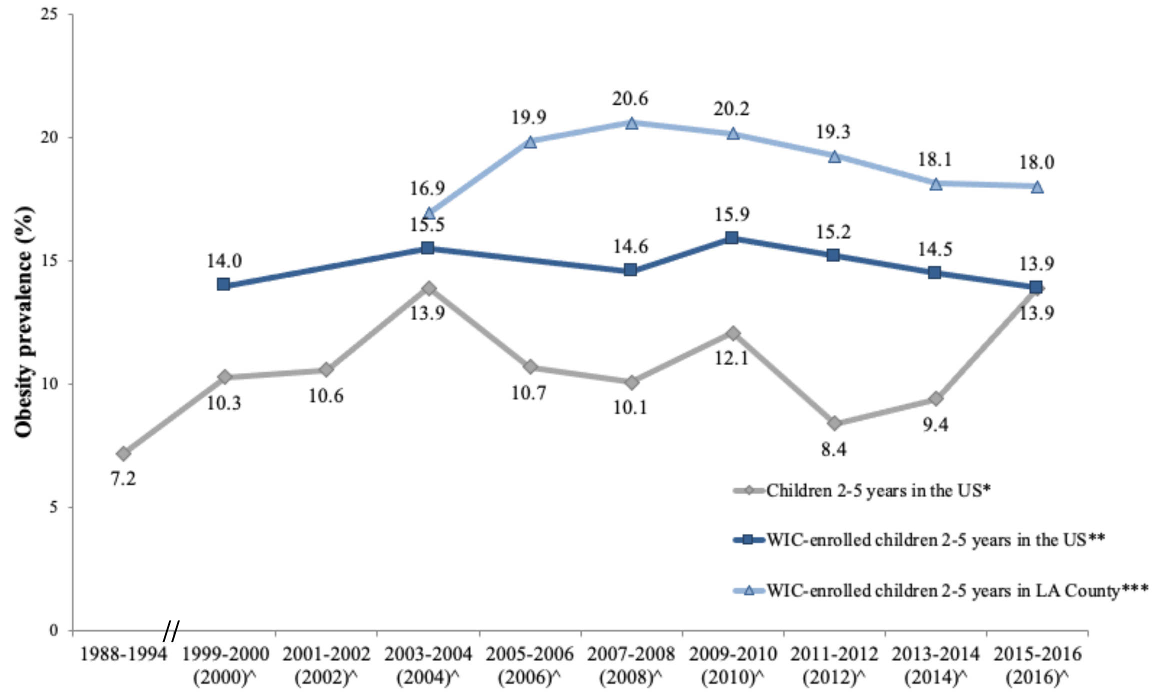
As childhood obesity is difficult to reverse through interventions,[11-13] early prevention is key. A Cochrane review of randomized control trials (RCTs) for preventing obesity in children found that diet and PA interventions combined may slightly reduce BMI (-0.07 kg/m^2; 95% confidence interval (CI): -0.14, -0.01) and standardized BMI (zBMI; -0.11; 95% CI: -0.21, 0.01) in children aged 0 to 5 years. For children aged 6 to 12 years, PA interventions may reduce BMI by 0.10 kg/m^2 (95% CI: 0.05, 0.14), but had little or no effect on zBMI.[11] However, the effects of interventions that combine both diet and PA on adiposity reduction among children 6 to 12 years and children 13 to 18 years remain unclear.[11]

Trends in early childhood obesity

Over the past three decades, obesity among preschool-aged children 2 to 5 years in the US has increased dramatically, almost doubling between 1988-1994 and 2003-2004 from 7.2% to 13.9%, then significantly declining to 9.4% in 2013-2014.[1] Most recently, obesity has begun to increase again among preschool-aged children, reaching 13.9% in 2015-2016.[3] Children from low-income families, who are eligible for the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), are more likely to be obese than children from high income families.[49] Until recently, obesity prevalence among WIC-enrolled preschool-aged children has consistently been higher than that of the overall US population of preschool-aged children (**Figure 1.1**). In 2004, early childhood obesity prevalence was 15.5% among WIC-enrolled children compared to 13.9% in 2003-2004 for the overall US population, and 14.5% in 2014 compared to 9.4% in 2013-2014.[1,50] The most recent estimates available suggest that this disparity may have narrowed, with obesity prevalence reaching 13.9% in 2016 (2015-2016) for both the WIC-enrolled and the overall US population of preschool-aged children.[3,51]

In Los Angeles County (LAC), the most populous county in the US with an ethnically diverse population of over 10 million residents,[52] about half of all infants and children under 5 years (or over half a million children) are enrolled in WIC.[53] Obesity prevalence among WIC-enrolled children 2 to 5 years in LAC has consistently been higher than that of the overall US population of preschool-aged children and overall WIC-enrolled population of children 2 to 5 years in the US, reaching a peak of about 20% in 2008, and then decreasing to 18% in 2016 (see **Figure 1.1**).[54]

Figure 1.1. Trends in obesity prevalence for children 2 to 5 years



^Years corresponding to obesity prevalence for WIC-enrolled children
*Data from the National Health and Nutrition Examination Surveys (NHANES)
Sources: Ogden CL et al. JAMA. 2016 Jun 7;315(21):2292-9; Hales CM et al. NCHS Data Brief. 2017 Oct;(288):1-8.
**Data from the Pediatric Nutrition Surveillance System and the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC); Sources: CDC. MMWR Morb Mortal Wkly Rep. 2009 Jul 24;58(28):769-73; Pan L et al. MMWR Morb Mortal Wkly Rep. 2016 Nov 18;65(45):1256-1260; Pan et al. JAMA. 2019 Jun 18;321(23):2364-2366.
***Source: LA County WIC Data. Los Angeles County WIC Data. http://lawicdata.org/about/

More recent early childhood obesity prevalence estimates are currently unavailable at the national level. Childhood obesity prevalence estimates for US children are provided by the National Health and Nutrition Examination Survey (NHANES), which began in 1960. Since 1999, NHANES has been conducted by the National Center for Health Statistics continuously in 2-year cycles.[55] The 2017-2018 was the most recent cycle completed. The data set were to be released in early 2020. However, due to a lag time between data collection, cleaning, and analysis, childhood obesity prevalence estimates from the 2017-2018 cycle have yet to be released. NHANES provides unbiased demographic-specific obesity prevalence estimates for preschool-aged children in the US at the national-level. NHANES was not designed to be able to provide estimates at smaller levels of geography, such as state-level estimations. Early childhood obesity prevalence estimates for WIC-enrolled children were provided by the Pediatric Nutrition Surveillance System (PedNSS) for 1973-2011.[55] Beginning in 2012, the estimates are provided by the WIC Participant and Program Characteristics (WIC PC) survey, a biennial census of WIC participants nationwide conducted by the United States Department of Agriculture (USDA) in April of even years.[56] The most recent prevalence estimates from 2016, available at the national and state level, were released in June of 2019.[51]

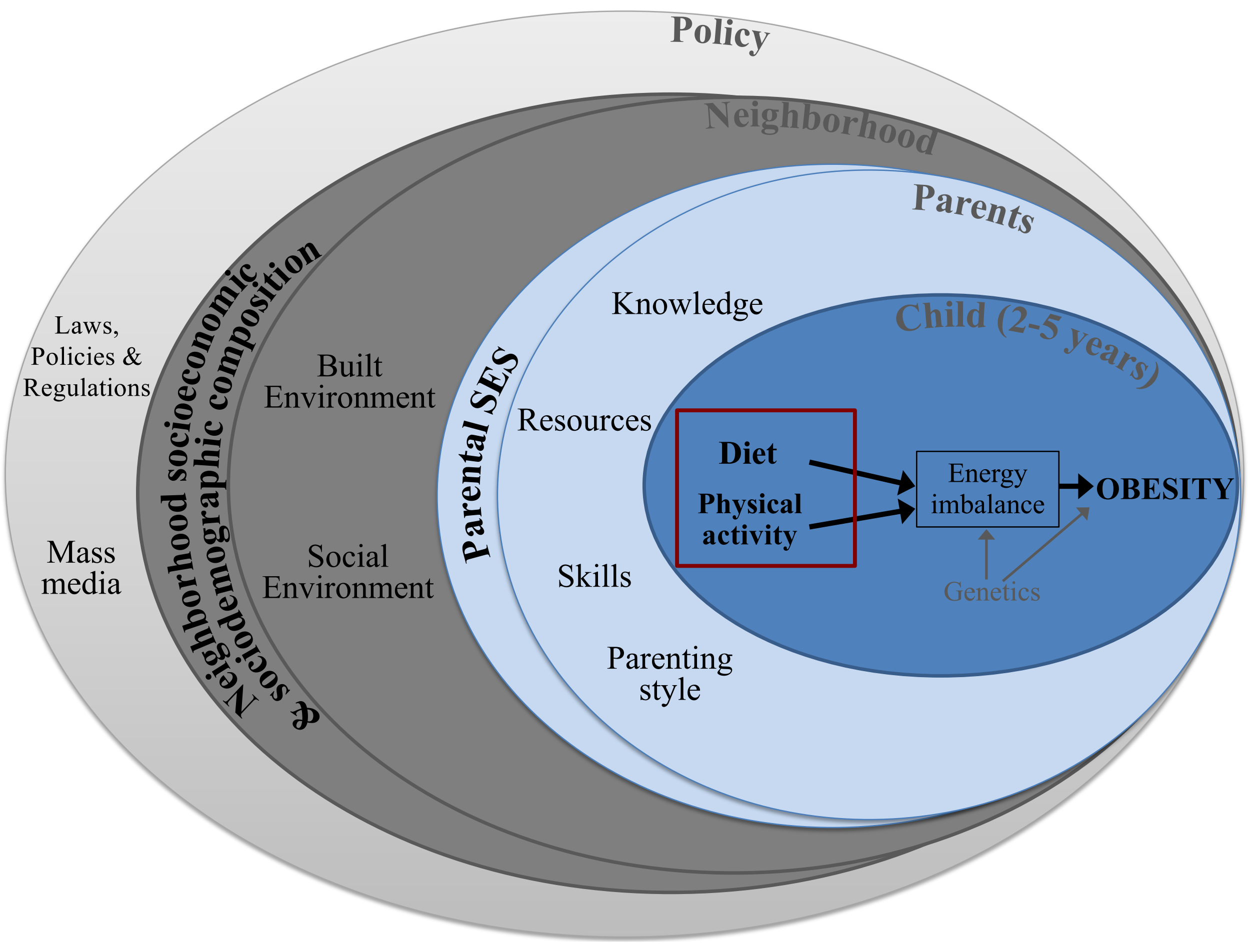
In LAC, data on obesity prevalence for WIC-enrolled children are available at the county level, as well as smaller geographic levels including census tracts and ZIP Codes.[54] The LA County WIC Data Mining Research Partnership between Heluna Health, formerly, Public Health Foundation Enterprises (PHFE) WIC, and First 5 LA provides access to a unique administrative database that contains information on WIC participant records for the duration that participants' receive WIC services (from the prenatal period through the child's fifth birthday) for every year since 2003.[54] LAC is the only county in the US that is able to electronically aggregate and analyze WIC data across all WIC agencies in the county.[54] Obesity prevalence estimates for 3- and 4-year old WIC-enrolled children are publicly available up to 2018.[57]

Causes and risk factors of obesity

The physiological and proximal cause of obesity is energy imbalance resulting from energy intake that exceeds energy expenditure. Excessive energy intake leads to the excessive accumulation of adipose tissue, or obesity. Energy intake is influenced by diet, whereas levels of physical activity (PA) largely determine energy expenditure.[9] The regulation of body weight and adiposity is a complex process that involves genetic, endocrine, behavioral, psychosocial, and environmental factors.[10]

Given the complexity of risk factors involved, the socioecological model[14] provides a framework for investigating the independent and synergistic contributions of various risk factors at different levels of the socioecological model to the development of early childhood obesity (**Figure 1.2**).

Figure 1.2. Conceptual model of risk factors related to the development of early childhood obesity



Individual-level characteristics

The primary and most proximate determinant of early childhood obesity occurs at the individual level of the child. Imbalance between energy intake and expenditure ultimately determines a child's obesity status, and is a direct result of the child's diet and levels of activity, although this relationship can be modified by genetics. Like height, weight growth can be inherited from parents with estimates on the heritability of excess body mass/fat clustering around 25 to 40%,[58] and heritability of obesity ranging anywhere from 20 to 80%.[59] Genetics, however, cannot explain the dramatic increase seen in obesity prevalence over the past few decades, both around the world and within the US.[60,61] Social determinants of health are the social and economic conditions that influence individual and group differences in health status.[62] Taking a "social determinants of health" approach, the non-biological social determinants of obesity will be discussed in the following section.

Parent-level influences

Parents play an important role in shaping a child's risk of early childhood obesity. As children transition from consuming a single food (breast milk or formula) to consuming a variety of solid foods, they begin to develop their own food habits, which are shaped by foods introduced to them by their parents,[63-65] as well as parents' feeding styles.[66,67] Parents also shape the home food environment by, for example, determining the availability of foods at home, portion sizes, and mealtime structure, which can either encourage or discourage healthy eating.[68] Health-related behaviors, including behaviors related to PA, track from childhood to adulthood,[69] indicating the importance of early and ongoing opportunities for PA. Parents can influence their child's development of healthy habits related to PA by exercising with their child (such as taking them to parks to play or going for family walks),[70] and providing encouragement and opportunities for PA (such as enrolling children in organized sports, attending child's sporting events, and providing transportation to PA events).[70,71] Parenting style has also been associated with children's levels of PA.[72] Parents' own behaviors related to diet and PA are also important. Consistent with Social Cognitive Theory,[73] children learn through observation and model the eating and PA behaviors of their parents.[70,71,74]

Parental socioeconomic status (SES) largely shapes the influence that parents have on their child's obesity-related behaviors. SES is an underlying dimension of social stratification and social ordering, made up of a combination of class components such as income, occupational status, and educational attainment.[75,76] Those with higher levels of SES have greater access to assets (or flexible resources), such as knowledge, money, and power, that can be used to avoid risks and adopt protective strategies to secure good health.[77] In the US, systemic racism was born of the economic advantages of slavery for whites when the nation was founded, and since that time, major institutions in the US have been pervaded by racial stereotypes, ideas, emotions, and practices, reproducing over time the socioeconomic conditions that reinforce systemic racism.[78] As a result, in the US, tightly interwoven with SES is a person's race/ethnicity, which, via systemic racism, leads to higher SES attainment and greater access to flexible resources for white persons, the "dominating racial group", compared to minority groups.[79] Access to knowledge, skills, and financial resources may be constrained for parents with lower levels of SES— resources that are needed to support the development of healthy obesity-related behaviors in young children. Parental SES also determines the neighborhoods that families can afford to live in, which further contextualize the development of behaviors important for early childhood obesity.

Neighborhood-level influences

The neighborhood built and social environments provide the context that can enable healthy behaviors or hinder them. Broadly defined, the built environment encompasses human-made or human-modified aspects of a person's surroundings. The neighborhood built environment affects energy balance by presenting opportunities or barriers to access healthy foods and opportunities or barriers for PA. A built environment with limited access to healthy food, convenient access to unhealthy energy dense foods, and few opportunities for physical activity is often referred to as obesogenic.[80,81] The "obesogenicity" of an environment is defined as "the sum of influences that the surroundings, opportunities, or conditions of life have on promoting obesity in individuals or populations."[17]

Neighborhoods have varying levels of access to, and density of, different types of food outlets such as supermarkets, grocery stores, convenience stores, fast-food establishments, full-service restaurants, and limited-service restaurants.[82] Supermarkets offer a large variety of healthy foods, and other types of food stores, such as convenience stores, are assumed to carry a larger proportion of high-calorie foods.[83,84] Fast-food restaurants, which often have limited and unhealthy food options, serve food that are higher in calories, fats, and carbohydrates compared to foods prepared at home.[85-87] Since the food choices that people make are limited to what is available to them, and convenience is an important predictor of food habits,[88] individuals living in areas with greater access to unhealthy food options may be more likely to adopt an energy-dense diet.[84] The neighborhood food environment shapes the proximity, accessibility, variety, and quality of healthy food options that parents, and therefore children have, thereby facilitating or constraining individual and family food choices.

The neighborhood PA environment, characterized by the neighborhood's physical design, land use patterns, transportation systems, and access to recreational facilities such as parks and playgrounds can influence an individual's likelihood of engaging in PA.[89-92] Neighborhoods that provide a range of local facilities, such as shops and food establishments, within walking or cycling distance, and supportive infrastructure like well-maintained sidewalks are often referred to as "walkable neighborhoods".[89] Neighborhoods with pedestrian-friendly walkways, greater access to public transit stops, high street connectivity, and mixed-used development can encourage active lifestyles, whereas neighborhoods lacking these characteristics can lead to increased reliance on cars, more heavily trafficked roads, and decreased opportunities for PA.[90,91,93] Access to recreational facilities, both in terms of distance and density, as well as the variety and quality of these facilities, is another important aspect of the built environment that contextualizes individual PA behaviors.[82] For young children, access to safe areas to play, playground density, as well as parental perceptions of playground safety, may be important factors influencing PA among young children.[94-96]

The neighborhood social environment may influence food consumption and PA through neighborhood social capital, collective efficacy, and social norms. Neighborhood social capital—the quality and quantity of social resources in a community,[97] and collective efficacy— the mutual trust between neighborhood residents who share beliefs and expectations that enable them to collectively work together to intervene, when necessary, for the good of the community,[98,99] strongly predict neighborhood crime.[98,100] Increased crime may discourage residents from spending time outdoors and constrain opportunities for children to play and spend time outdoors. Neighborhood collective efficacy may determine how neighborhoods deal with issues that affect the community. For example, neighborhood residents may work together to maintain and improve local environments, such as recreational resources. Social norms, defined as explicit or implicit rules that guide, regulate, proscribe, and prescribe social behavior in particular contexts,[101] may influence behaviors related to eating and PA.[82,102,103] Social norms in a neighborhood are shaped by the sociodemographic characteristics of the neighborhood's residents. The greater the concentration of like-minded people, the stronger the norms, and the greater the exposure of residents to these norms.[104] For example, the use of a common space, such as a park, by a given subpopulation within a neighborhood is likely to be influenced by local norms and whether or not residents feel "out of place" among the other users at a given time of day. These social environment factors can affect obesity by either supporting or discouraging behaviors related to obesity.

The neighborhood socioeconomic and sociodemographic composition largely shapes the characteristics of a neighborhood's built and social environment. Neighborhoods with higher concentrations of lower SES residents and minority residents are likelier to have fewer built and social resources compared to neighborhoods with higher concentrations of higher SES residents, which can restrict access to healthy food and opportunities for PA, while also impacting social norms around obesity-related behaviors.[91] The higher the concentration of lower SES residents and minority residents in a neighborhood, the less disposable income residents of that neighborhood have to support local shops, services, and restaurants. As such, the availability and types of businesses, including food outlets and for-profit recreational facilities are often related to neighborhood sociodemographic factors.[105-107] In the US, lower income and racial/ethnic minority neighborhoods tend to have fewer supermarkets, fruit and vegetable markets, parks, sports facilities, and bike paths, and more fast food restaurants.[107-111] According to social disorganization theory, neighborhood characteristics such as poverty, ethnic heterogeneity, high residential turnover rates, low homeownership rates, and concentration of recent immigrants may make it more difficult for residents to establish social ties, as well as build social capital and collective efficacy,[104,112] and these neighborhoods are likelier to suffer from higher crimes rates.[98,113,114]

Policy-level influences

Laws, policies, and regulations, as well as mass media may intentionally or inadvertently contribute to obesogenic environments in neighborhoods and homes. Legislation around, for example, food subsidies, agriculture, trade, urban planning, or transport, may contribute to obesogenic environments in various ways. Laws, policies, and regulations set at various levels of government can trickle down through the layers of the socioecological model to influence parents' and children's obesity-related choices and behaviors. For example, in the US, government regulations affect food prices through farm policies that subsidise fat and sugar production, and keep fruit and vegetable prices high.[115,116] Food prices consistently influence individuals' food purchases, and households with limited resources may be more likely to choose cheaper energy-dense, high-sugar, and high-fat foods that provide calories at the lowest cost.[115,117] Urban planning regulations that promote single, rather than mixed, land use in cities may make it more difficult for residents to stay active, whereas planning that prioritizes active or public transport over vehicle use can encourage PA.[90,116] Mass media can influence society's attitudes, beliefs, values, and norms around food preference and social desirability of exercise, by not only reflecting and reinforcing the current culture, but also shaping it, through advertising, marketing, or mass media campaigns.[118,119]

Addressing the complexities of early childhood obesity

Given the complexity of risk factors involved in the development of early childhood obesity, and the synergies between the different risk factors, interventions focused solely on individuals to prevent or reduce childhood obesity have demonstrated limited success.[11-13] A multi-pronged approach guided by a socioecological framework[14] to address the myriad of factors may be necessary.[15,16]

Past and current approaches obesity control

Over the past few decades, approaches to health promotion have progressed from individual-based approaches to community- and population-based approaches that incorporate socioecological perspectives.[120] To address the multiple factors and levels of influence on behaviors, community-based programs began to implement multicomponent, multilevel interventions to target change among individuals, groups, and organizations, while also attempting to incorporate strategies to create policy and environmental changes.[121] Multicomponent interventions refer to interventions that incorporate more than one intervention strategy to achieve an improved health outcome, where an intervention strategy is defined as a plan of action that describes a method for achieving project objectives and producing defined outcomes. Multilevel interventions refer to interventions that address more than one level of the socioecologic model,[14] for example, an intervention that includes a component to teach parents about recommended levels of PA for their child (parent level) as well as a component that provides the family with access to safe, high-quality space for play, exercise, and recreation through a joint use agreement with a local school (neighborhood level).

The first generation of community-based prevention programs were conceptualized in the 1960s with the intention of reducing growing rates of cardiovascular disease (CVD) at the population-level, and focused on identifying and treating high-risk individuals[122,123] using a "high-risk" approach largely based on the medical model.[124] The strategy evolved during the 1970s and 1980s after the importance of behavioral influences on health were identified from the Framingham Heart Study.[125] Rather than targeting high-risk individuals, the "population" approach proposed by Rose to target the entire population— regardless of risk factor, disease status, or need, has proved to be more effective in reducing rates of disease at the population level.[126]

During the 1970s, multicomponent and multilevel interventions designed to address behavioral and social risk factors for CVD were implemented in North Karelia, Finland, and three small communities near Stanford University.[122,127] These two studies provided the foundation for three National Heart, Lung, and Blood Institute (NHLBI) research and demonstration projects in the 1980s—the Stanford Five-City Project, the Minnesota Heart Health Program, and the Pawtucket Heart Health Program, which were rigorously designed, well-funded experiments aimed to test the effectiveness of comprehensive, community-wide interventions in large, diverse American populations.[123,128] During the evaluation phase of these community-level interventions, the researchers found the intervention effects to be in the expected direction, but not statistically significant, and emphasized the analytic challenges of evaluating community-based prevention programs due to smaller than expected net differences.[128] However, by the end of the 1990s, health promotion and disease prevention interventions shifted their focus from individual lifestyle behaviors to community-level place-based strategies.[129] This community-based paradigm has been applied to the control of obesity.[18-20]

Over the last decade, more than $2 billion was pledged by CDC, the Robert Wood Johnson Foundation (RWJF), and The California Endowment to target and help reverse the obesity epidemic through community programs, policies, and interventions.[18-20] Obesity is being addressed through community-level place-based health promotion initiatives that are multilevel, multicomponent, and implemented through multiple sectors and settings.[21-23] These place-based "whole of neighborhood" interventions seek to reduce obesity prevalence for entire populations.

Building on Rose's "population" approach, Frohlich and Potvin have argued that disparities in health may be exacerbated by interventions that target the entire population, and propose the "vulnerable population" approach to target under-resourced populations, rather than the entire population, to reduce health disparities.[130] Identification of these populations is paramount to reducing disparities in obesity between communities, in addition to allowing obesity control programs to be tailored to the specific needs of communities, while leveraging on communities' existing resources. Programs tailored to the specific needs of the communities they serve are more likely to be effective than broadly based programs that target entire populations or programs delivered at higher scales of geography, such as programs delivered at the county or state level.[131-133]

In wealthy countries, such as the US, a number of studies have reported higher rates of obesity in under-resourced neighborhoods.[107,134-136] For example, living in neighborhoods with higher levels of poverty and lower levels of education has been associated with increased risk of childhood obesity after considering the effects of relevant individual-level risk factors.[137] Risk of obesity is also higher in communities of color, with obesity more prevalent among Hispanic (22.0%) and non-Hispanic Black (20.8%) children compared with their white (15.9%) and Asian (12.8%) peers.[138]

Many childhood obesity efforts, such as the RWJF's Healthy Kids, Healthy Communities— a national program to implement healthy eating and active living initiatives— focus their efforts on lower-income communities and communities of color, where the risk of obesity is greatest.[139] However, to our knowledge, no published studies have examined how resources from various funding agencies are distributed to address childhood obesity among communities at risk. It is unclear if resources are reaching all communities with need or what factors determine the allocation of scarce resources. Furthermore, current data limitations may thwart the efforts of funders and policymakers to identify communities with the highest prevalence of early childhood obesity.

Identifying communities most burdened by obesity

To reduce disparities in obesity prevalence, communities that have the highest prevalences of early childhood obesity should be targeted. However, to identify these communities, timely data on early childhood obesity prevalence is needed, in addition to estimates at the appropriate level of geography that better align with administrative planning areas. Obesity prevalence estimates are currently available from surveillance data that are collected by national and state health surveys or surveillance systems. However, there are issues with the regularity, timing, and granularity of these publicly available sources of data. National and state health surveys and surveillance systems have significant lag times between data collection and data availability, and often have gaps between waves of data collection.

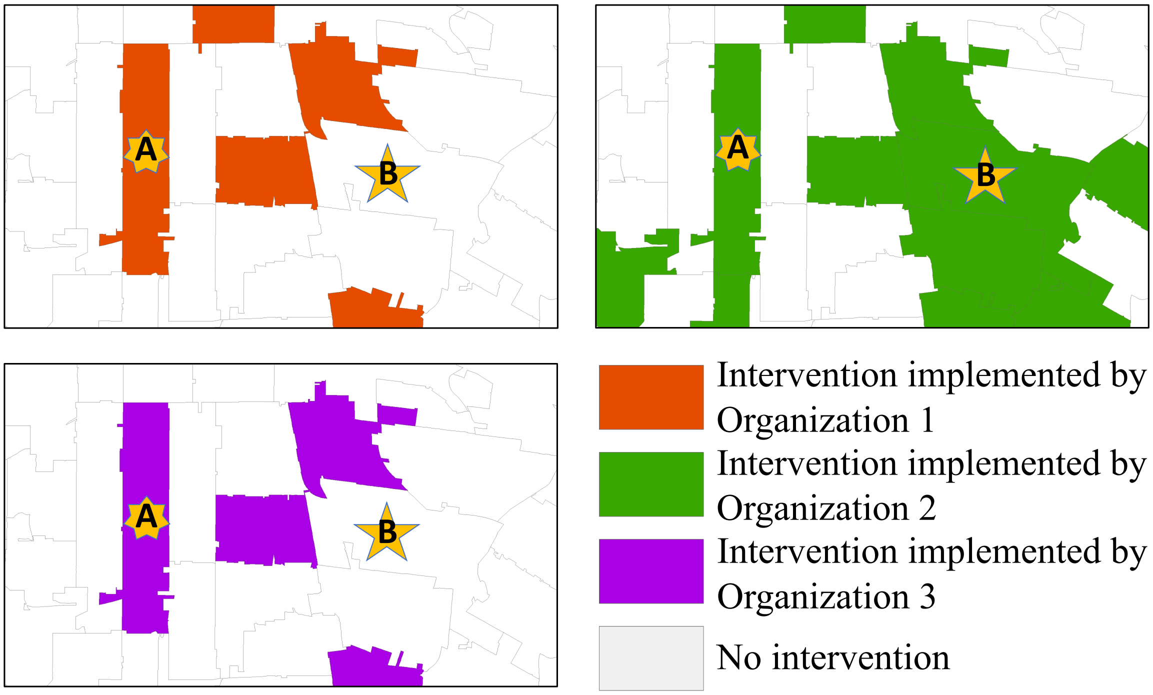
At the national and state levels, data from surveillance systems and surveys that can be used to calculate obesity prevalence estimates include NHANES, WIC PC, the Behavioral Risk Factor Surveillance System (BRFSS), the California Health Interview Survey (CHIS), the National Health Interview Survey (NHIS), the National Survey of Children's Health (NSCH), the North Carolina Child Health Assessment and Monitoring Program (NC-CHAMP), the Panel Study of Income Dynamics (PSID), and the Panel Study of Income Dynamics Child Development Supplement (PSID/CDS).[55,56] Among publicly available data sources, obesity data for children 2 to 5 years, our population of interest, are only available from NHANES, WIC PC, CHIS, NC-CHAMP, PSID, and PSID/CDS. In addition to the lag times of data availability, as discussed, obesity data for early childhood is only available at the national or state level, with the exception of CHIS that provides estimates at the county level.[140] The geographic level of these survey and surveillance data are insufficient as they do not align with the geography of local place-based obesity initiatives. However, in addition to these national and state level surveillance systems, many cities, towns, and regions in the US have their own surveillance data available at smaller levels of geography to better monitor, evaluate, and improve the health of the populations they directly serve. For example, we previously described the LA County WIC Data Mining Project database, which can be leveraged to estimate obesity prevalence among WIC-enrolled children in LAC across census tracts or ZIP Codes.[54] However, as with most surveillance systems, the lag times between data collection and data availability continue to prevent funders and policymakers from accessing quick, up-to-date estimates of obesity prevalence across communities.

Novel approaches to intervention research

Sociodemographic data are routinely collected by the US Census Bureau, and the data are easily accessible to the public by various levels of geography. As discussed previously, neighborhoods and the sociodemographic characteristics of neighborhoods shape obesity risk for individuals living in those neighborhoods.[134-136,141] However, as is the case with surveillance data, there is a lag time between data collection and availability of sociodemographic data. Can we use existing data, that is, the most recent estimates of both obesity prevalence and sociodemographic data available, to predict future prevalence of obesity? Machine learning is the study of computer algorithms that can be automated to continuously improve and learn from past experiences.[142] Combining the predictive capabilities of machine learning algorithms with the most recent estimates of obesity prevalence data available, and the availability of sociodemographic data at various levels of geography may provide an opportunity to build a statistical model that could allow funders and policymakers to quickly and accurately identify neighborhoods where obesity risk is highest. If we are able to build an accurate model that uses the most recent obesity prevalence and sociodemographic data available to predict future obesity rates at the neighborhood level, we would be able to overcome the current issue of lag times between data collection and data availability. In addition to identifying neighborhoods with the greatest needs for place-based obesity prevention efforts, funders and policymakers also need to know which intervention components are most effective at reducing community-level obesity.

Place-based interventions, especially interventions that are multilevel and multicomponent, are difficult to evaluate. The "gold standard" for evaluating community-level interventions is the cluster randomized trial (CRT), where communities, rather than individuals, are randomized into an intervention or a comparison group. However, CRTs are often expensive, challenging to implement, and it is unlikely for communities to agree to be randomly assigned to a comparison group. It is also impractical and often inappropriate to implement a "one size fits all" intervention for all communities given that communities have different risk factors and risk conditions. Due to these challenges, many place-based interventions are not CRTs. Another obstacle to evaluating place-based interventions is how best to correctly measure or quantify the intervention or interventions that a given community received. With the influx of funding to address obesity at the community level, many communities now receive a variety of interventions simultaneously, as various organizations implement interventions without coordination. As an illustration, in **Figure 1.3**, we can see that Community A received an intervention implemented by Organization 1, Organization 2, and Organization 3, respectively. However, Community B only received an intervention implemented by Organization 2. Without identifying all of the interventions that a community received at a given time, we are unable to understand how the various place-based interventions contributed to any changes in obesity prevalence seen at the community-level.

Figure 1.3. Illustration of the simultaneous exposure of community-level interventions



Studies have shown that place-based interventions show some promise for preventing obesity, leading to modest reductions in population weight gain,[21,22,24] albeit the evidence has been mixed.[23,25] A review of place-based interventions targeting children or adolescents found that six of the eight reviewed studies reported a significant improvement in at least one measure of adiposity that could be attributed to the intervention, with a pooled mean reduction in Body Mass Index z-score (zBMI) of 0.09.[21] Another review that included 51 school-based interventions found that interventions to improve weight status in preschool-aged children found small magnitudes of effect sizes, and led to smaller BMI increase over time among children who received the intervention relative to the comparison group.[24] Conversely, a review of multilevel and multicomponent obesity-related interventions found that only three of the eight studies included in the review reported significant reductions in obesity.[23]

Based on the published literature, place-based interventions seem to have achieved some success in reducing BMI. The studies included in the reviews, however, did not account for scenarios where communities received various interventions simultaneously, which would overestimate the effects of the interventions evaluated. Furthermore, the majority of studies included in the reviews had a control or comparison group. In a "real world" scenario where communities receive various interventions, simultaneously, are we able to find an appropriate comparison group? And in doing so, can we identify which types of interventions have contributed the most to declines in obesity prevalence seen at the community-level? In LAC, we showed that obesity prevalence among WIC-enrolled children aged 2 to 5 has decreased in LAC since peaking at about 20% in 2008. While surveillance data can provide information on prevalence and trends in obesity, it cannot explain the mechanisms associated with changes in obesity prevalence, and it is currently unclear which types of interventions are most effective at reducing community-level obesity rates.[143]

To address the gaps in the literature we discussed in this chapter, we use 3 different datasets to:

(1) Examine the distribution of community-level interventions addressing early childhood obesity in Los Angeles County to understand how resources for obesity prevention programs are allocated across communities;

(2) Determine whether existing sociodemographic and obesity prevalence estimates can be used to identify communities most burdened by obesity in a timely manner; and

(3) Examine the contribution of place-based interventions to declines in neighborhood-level rates of early childhood obesity, and identify the types of interventions that produce the greatest reductions in neighborhood-level early childhood obesity.

The 3 datasets used in this dissertation are described in the next chapter.

CHAPTER 3: Data Sources

In this dissertation, we used the following 3 datasets, described below:

I. LA County WIC Data Mining Project;

II. Early Childhood Obesity Systems Science Study (ECOSyS) Intervention Database;

III. The Decennial Census and the American Community Survey (ACS).

I. LA County WIC Data Mining Project

The Special Supplemental Nutrition Program for Women, Infants and Children (WIC) is a federally funded nutrition program for low-income pregnant, breastfeeding, and postpartum women, infants and children under 5 years of age who are at risk for poor nutrition.[144] WIC is a short-term program that provides eligible WIC participants with supplemental foods, health care referrals, and nutrition education for the duration of a certification period, usually 6 months to a year.[145] Once the certification period ends, WIC participants must reapply.

In Los Angeles County (LAC), WIC serves about half of all children under 5 years of age, or over half a million children.[53] WIC recipients are served by 7 WIC providers in LAC, the largest of which is Heluna Health, formerly, Public Health Foundation Enterprises (PHFE) WIC.[146] In 2002, the LA County WIC Data Mining Project was initiated by PHFE WIC (now Heluna Health), through agreements with the State of California, and funded by First 5 LA (F5LA) through a research partnership to collaborate with the six other local agency WIC Programs in LAC to collect and analyze WIC data.[146,147] F5LA is an independent public agency, created by voters in 1998 to invest LAC's allocation of funds from California's voter-approved Proposition 10 tax revenues, whose goal is to support the safe and healthy development of young children.[148] LAC the only county that is able to electronically aggregate and analyze WIC data across all of the WIC agencies in a county.[147] The primary goal of the LA County WIC Data Mining Project is to provide comprehensive data to examine and address the needs of the low-income WIC population in LAC.[146]

The LA County WIC Data Mining Project gathers and maintains data routinely collected on every WIC participant in all seven local WIC agencies in LAC.[146] The database, with data on about 3.5 million WIC participants since 2003, includes data on child's sociodemographics, health information, home addresses, as well as data on age, height, weight for the duration participants' receive WIC services (from the prenatal period through the child's fifth birthday).[146] These data are regularly updated due to federal regulations for WIC recertification that require children's height and weight to be measured at least once a year.[145] WIC staff are trained to use a standardized protocol to measure child's height using wall-mounted stadiometers (Model PE-WM-60-76; Prospective 43 Enterprises, Portage, MI), and weight using calibrated beam scales (Health-O-Meter 402LB; Prospective Enterprises, Portage, MI) during clinic visits, which have been shown to have high accuracy.[149] Aside from measurements taken during clinic visits, about 20% of height and weight data are taken from pediatric provider records, which can be used if a child visits a health care provider within 60 days of the WIC recertification appointment.[149]

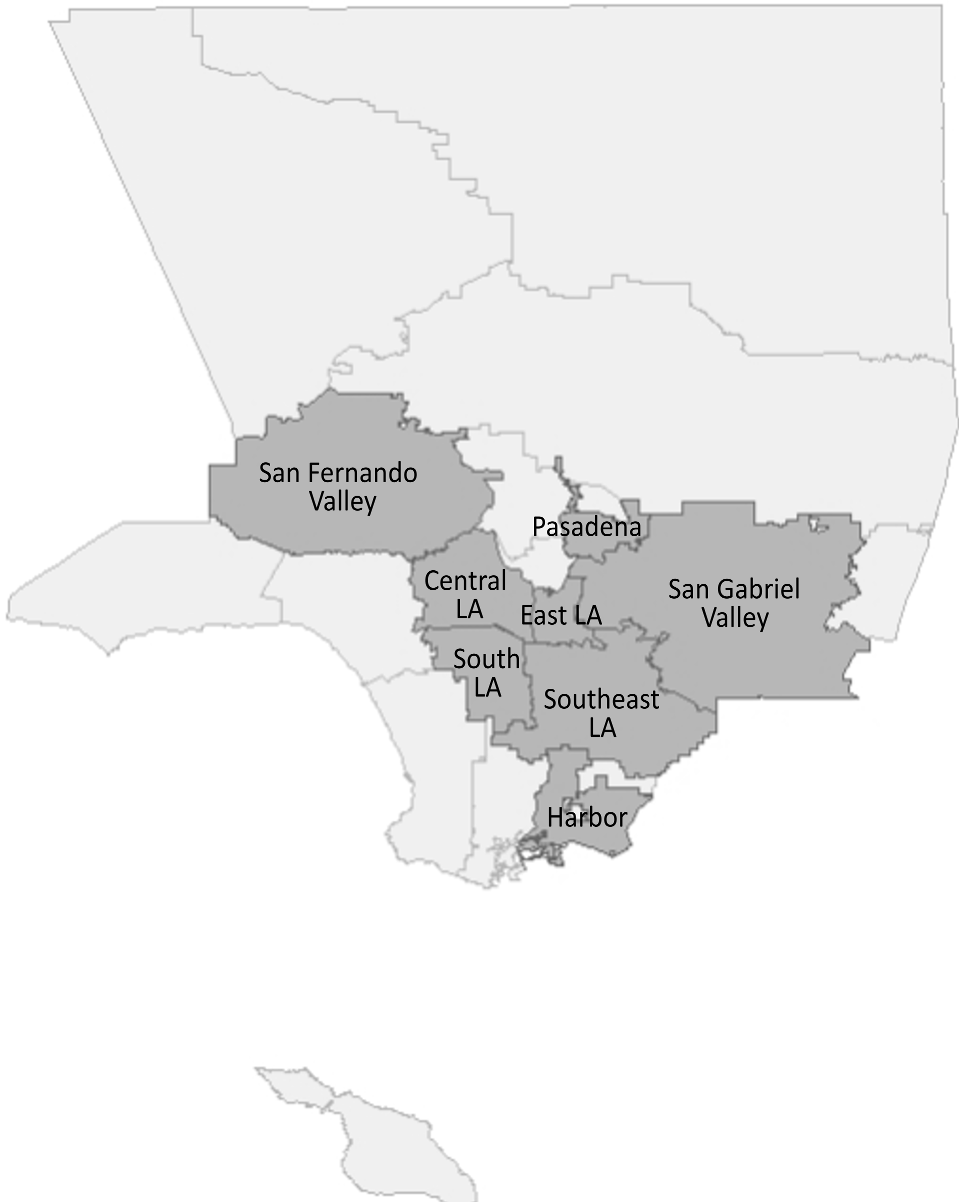
II. Early Childhood Obesity Systems Science Study (ECOSyS) Intervention Database

The Early Childhood Obesity Systems Science Study (ECOSyS) is a partnership that was established in 2013 among UCLA, PHFE-WIC (now Heluna Health), the Los Angeles County Department of Public Health, the University of Washington, the University of California at

Berkeley, the Samuels Center, F5LA, and Kaiser Permanente, and supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD, to pioneer the use of causal inference and systems science methods for evaluating community interventions.[150] The primary aim of ECOSyS is to investigate the independent and combined effects of obesity-related public policies and community interventions on obesity in preschool-aged children in LAC.

The ECOSyS Intervention Database was created to identify and characterize obesity-related interventions implemented in communities in LAC since 2003, with a focus on major initiatives.[148,151-153] Relevant data were obtained from organizations that funded major obesity initiatives in LAC, reports and websites of both grantor and grantee organizations, Tax 990 forms, and interviews with program staff of WIC clinics that sought external funding for interventions beyond those that are mandated by WIC. Most tax-exempt organizations (such as nonprofit organizations) are required to file Tax 990 forms to the US Internal Revenue Service (IRS), which provide an overview of the organization's activities, governance, and detailed financial information. Interviews with WIC clinic program staff were conducted from WIC clinics that operated within 8 Regions in LAC, where the majority of WIC families reside (**Figure 3.1**). The smallest geographic unit for which intervention data was able to be extracted was at the ZIP Code level.

Figure 3.1. Regions in Los Angeles County where interviews with WIC clinic program staff were conducted



The ECOSyS Intervention Database characterizes obesity-related interventions by year, implementing organization, funding organization, start and end dates, ZIP Code(s) reached or targeted, program names, project names, and specific intervention strategies. A program is defined as a long-term managed portfolio of multiple projects to produce outcomes. Projects are short-term and designed to deliver a specified output and achieve outcomes within a specified time period and location, and may use various intervention strategies. An intervention strategy is defined as a plan of action that describes a method for achieving project objectives and producing defined outcomes.

Each intervention program may include one or several projects, and each project may use one or more (up to 10) intervention strategies (**Figure 3.2**). Intervention strategies used by each project were determined using a modified typology that identified 10 strategies, expanding on the 9 strategies identified by the original typology,[150,154] broadly categorized as macro-level or micro-level (**Table 3.1**). A macro-level strategy is defined as an intervention strategy that does not directly target individuals but may affect the larger community. The 4 macro-level intervention strategies identified were: (i) government policies, (ii) public institutional policies, (iii) infrastructure investments, (iv) business practices. A micro-level strategy is defined as an intervention strategy that directly targets individuals. The 6 micro-level intervention strategies identified were: (i) group education, (ii) counseling, (iii) health communication & social marketing, (iv) home visitation, (v) screening & referral, and (vi) staff training. In the original typology of childhood obesity intervention strategies[150,154] staff training was grouped together with counseling. Each project that reaches or targets a ZIP Code contributes a minimum of 1 and a maximum of 10 intervention strategies to that ZIP Code in a given year.

Figure 3.2. Hierarchical classification of interventions addressing obesity

Program — Long-term managed portfolio of multiple projects to produce outcomes

➔ **Project** — Short-term and designed to deliver a specified output and achieve outcomes within a specified time period and location, and may use various intervention strategies

➔ **Intervention strategies** — A plan of action that describes a method for achieving project objectives and producing defined outcomes

Macro-level (*indirectly affects the larger community and obesity-related behaviors & practices*)**:**

- Government policies
- Public institutional policies
- Infrastructure investments
- Business practices

Micro-level (*directly targets a specific population and obesity-related behaviors & practices*)**:**

- Group education
- Counseling
- Health communication & social marketing
- Home visitation
- Screening & referral
- Staff training

Table 3.1. Typology and description of intervention strategies

MACRO-LEVEL INTERVENTION STRATEGIES[1]
1) Government policies: National, state or local policies (e.g., principles, rules, guidelines, legislation) that aim to influence the accessibility of healthy and unhealthy foods, increase opportunities for physical activity, improve healthcare access, or promote breastfeeding. *Examples:* Food subsidies to support locally grown foods; food taxes on sugar-sweetened beverages; zoning laws to limit fast food operations; regulation of food marketing practices targeting children; tax breaks to businesses that provide on-site recreational facilities for exercise; health insurance for low-income children; longer maternity leave
2) Public institutional policies: Policies by public institutions such as county governments, school districts, Head Start childcare programs, and healthcare facilities that aim to increase the accessibility of healthy (vs. unhealthy) foods, increase opportunities for physical activity, or promote breastfeeding. *Examples:* Nutritional guidelines for food procurement and foods served; mandatory physical education for students; schools allowing their facilities to be used by residents during weekends (joint-use agreements); baby friendly hospital policies
3) Infrastructure investments: Efforts to change the physical environment to promote healthy eating and active living. *Examples:* Walkable neighborhoods; parks; establishment of healthy food venues (e.g., farmers' markets, supermarkets)
4) Business Practices: Practices by the private sector that influence consumer choice and decision-making. *Examples:* Product placement in a grocery store; restaurant procurement of locally grown foods; menu changes; menu-labeling
MICRO-LEVEL INTERVENTION STRATEGIES[2]
1) Group education: An intervention that involves imparting knowledge and/or skills to a group of individuals, including breastfeeding workshops; nutrition education, exercise and parenting classes; and cooking demonstrations.
2) Counseling: Interactions with the child and/or child's parent/caregiver by a trained counselor or para- professional with the goal of changing food consumption patterns of the child, parenting style, or parenting practices.
3) Health communication & social marketing: The use of communications strategies, consumer research, and/or marketing principles to promote health by influencing individual decisions that affect health.
4) Home visitation: A program that primarily delivers family-oriented services through home-visiting and may address parenting practices and child feeding practices.
5) Screening & referral: A program that screens for suboptimal growth (e.g., overweight/obesity) and/or inadequate nutrition, and refers the child to appropriate programs such as WIC.
6) Staff training*: Staff training, which was originally grouped together with the micro-level strategy 'group education' for which the training was provided.

Typology modified from Wang et al. 2018[150]

[1]Strategies that indirectly affect the larger community and obesity-related behaviors and practices

[2]Strategies that directly target individuals to modify obesity-related behaviors

*Originally grouped together with counseling[150]

III. The Decennial Census and the American Community Survey (ACS)

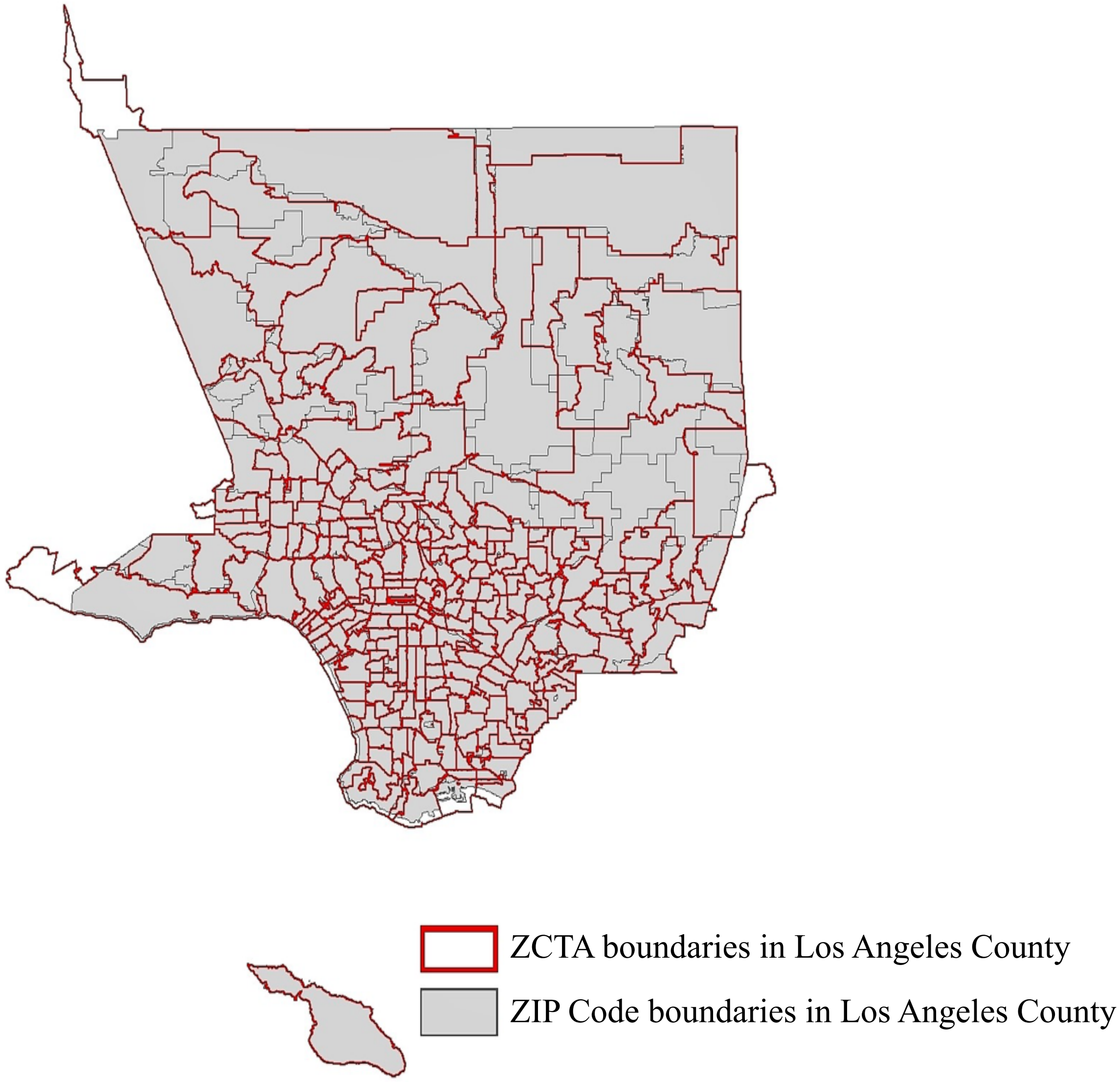
In the US, a census has been conducted every 10 years since 1790.[155] Between 1970 and 2000, the US Census Bureau used two different questionnaires— a short form and a long form. The short form census was designed to collect basic demographic and housing information (such as age, race, sex, relationship, and tenure), and the long form, sent to approximately 1 in 6 households, collected social, housing, and economic information (such as citizenship, educational attainment, disability status, employment status, income, and housing costs).[155]

Since 2005, in order to provide communities, businesses, and the public with detailed information more frequently, data that were historically collected once every 10 years by the decennial census long form, have been collected monthly (and released annually) through the American Community Survey (ACS).[155] The US Census Bureau contacts over 3.5 million households across the country to participate in the ACS every year.[155] At the level of ZIP Code Tabulation Areas (ZCTA), ACS 5-year estimates are available beginning 2006-2010.[156]

Beginning in 2000, the US Census Bureau created ZCTAs as generalized areal representations of United States Postal Service (USPS) ZIP (Zone Improvement Plan) Codes because ZIP Codes were so commonly used.[157] ZIP Codes are a system of postal codes assigned by USPS to a section of a street, a collection of streets, an establishment, structure, or group of post office boxes, for the delivery of mail.[158] ZIP Codes are not areal features, but a collection of mail delivery routes, and are periodically changed to support more efficient mail delivery.[157] Since ZIP Codes are not true geographic entities, it is not possible to precisely determine the extent of territory they cover.[158] Unlike ZIP Codes, ZCTAs are statistical geographic entities created by aggregations of census blocks that have the same predominant ZIP Code associated with the residential mailing addresses based on US Census Bureau's Master Address File.[157,159] However, ZCTAs do not precisely depict ZIP Code delivery areas, and do not include all ZIP Codes used for mail delivery.[157,159] For example, ZCTAs do not include ZIP Codes assigned to areas that are primarily nonresidential. Since 2010, large water bodies and large unpopulated land areas are also not assigned ZCTAs.[157] Though ZCTAs follow census block boundaries, ZCTAs are independent of all other statistical and governmental entities, and frequently cross the boundaries of other geographic boundaries, such as counties and states.[157]

A map comparing ZIP Code and ZCTA boundaries within LAC can be seen in **Figure 3.3**. Although ZIP Codes and ZCTAs share identical identifier codes, for example, 90001 or 90002, this does not guarantee that they share the same geographic boundary. However, cartographically, there is relatively little discernable difference between ZIP Codes and ZCTAs.[160]

Figure 3.3. Comparison of ZIP Code and ZCTA boundaries in Los Angeles County



Sources: ZIP Code boundaries based on parcels: Los Angeles County GIS Data Portal 2018[161]
ZCTA boundaries: US Census Bureau[162]

CHAPTER 4: Examining community-level interventions addressing early childhood obesity in Los Angeles County

Introduction

Obesity has been declared a global epidemic by the World Health Organization (WHO),[163] and childhood obesity is one of the most serious public health challenges of the 21st century.[4] Overweight and obese children are more likely to be obese as adults, and to develop non-communicable diseases, such as type 2 diabetes and cardiovascular disease, at an earlier age compared to normal weight children.[4]

Obesity in childhood starts very early in life. In the US, its prevalence among preschool-aged children (2 to 5 years) was 13.9% in 2015-2016,[3] which is a nearly 50% increase from the 9.4% who were obese in 2013-2014.[1] Studies suggest that in the US, children from low-income families are at higher risk for obesity.[164] In particular, data from the Special Supplemental Nutrition Program for Women, Infants and Children (WIC)— a federally funded nutrition program for low-income pregnant, breastfeeding, and postpartum women, infants and children under 5 years old who are at risk for poor nutrition, show that preschool-aged children participating in WIC have been disproportionately affected by obesity.[50,144] Since 2000, when obesity estimates for WIC-enrolled children became available, obesity prevalence among preschool-aged children enrolled in WIC has been consistently higher than national estimates for the same age group, peaking at 15.9% in 2010,[50] compared to 12.1% for the overall US population in 2009-2010.[1] For the first time since 2000, obesity prevalence for both preschool-aged children enrolled in WIC and nationally reached 13.9% in 2016 (2015-2016).[3,51]

Understanding factors related to the recent decrease in overall obesity prevalence, specifically among vulnerable populations, can have important public health implications.

Biologically, obesity is due to energy imbalance, that is, energy intake that exceeds energy expenditure. Energy intake is determined by diet, while energy expenditure is determined by physical activity and basal metabolic rate.[9] The regulation of body weight is a complex process that involves genetic, endocrine, behavioral, psychosocial, and environmental factors.[10] Consequently, addressing obesity has been very challenging and a multifactorial approach guided by a socioecological framework[14] to address this myriad of factors may be necessary.[15-17] Such an approach is supported by an expanding body of literature on the roles that social and physical environments play in influencing diet,[165,166] physical activity,[165,167] and even metabolism.[168,169]

A number of studies have reported higher rates of obesity in under-resourced neighborhoods.[134,170-172] In particular, living in neighborhoods with higher levels of poverty, lower levels of education, and a high proportion of non-Hispanic Black or Hispanic residents has been associated with increased child obesity risk after considering the effects of relevant individual-level risk factors.[137] In such neighborhoods, the lack of safe parks and recreational facilities, high crime rates, poor accessibility of quality fresh food, and ready availability of cheaper processed food may present barriers to active living and healthy eating.[134] Rates of obesity are also higher in communities of color,[138] with obesity more prevalent among non-Hispanic Black (20.8%) and Hispanic (22.0%) children compared with their white (15.9%) and Asian (12.8%) peers.[138] Consequently, health promotion and prevention efforts are now addressing these multilevel determinants of obesity by implementing place-based "whole of neighborhood interventions" to reduce obesity prevalence among entire populations, addressing environmental risk conditions, in addition to individual risk factors.[21,129,173]

Over the last decade, more than $2 billion was pledged by the Centers for Disease Control and Prevention, the Robert Wood Johnson Foundation, Kaiser Permanente, and The California Endowment to reverse the obesity epidemic through place-based initiatives.[18-20,152] Los Angeles County (LAC), the most populous county in the US with an ethnically diverse population of over 10 million residents,[52] has been a recipient of significant funding to tackle obesity in under-resourced communities. LAC has a large, ethnically diverse population of over 10 million residents,[52] with neighborhoods that vary greatly in sociodemographic profile. While it is home to Beverly Hills, the region of South LA has among the poorest neighborhoods in the country with low levels of education and a high proportion of minority residents. About half of residents living in South LA do not have a high school degree, over 40% of its residents live on a household income below $20,000, over 50% of its residents are Hispanic, and almost 40% are non-Hispanic Black.[174] Adjacent to South LA is Southeast LA where its sociodemographic profile suggest needs that are almost as dire.[174] Other under-resourced neighborhoods can be found throughout LAC, in the regions and communities of San Fernando Valley, San Gabriel Valley, Antelope Valley, Pomona Valley, and Harbor, where about 30% of residents do not have a high school degree, and 40-50% have a household income of less than $40,000.[174]

In LAC, large local initiatives to address obesity began in about 2005 when several government, private, and health systems organizations began implementing place-based obesity prevention efforts.[151,175-177] As considerable amounts of funding are being spent to tackle obesity, decisions have to be made regarding the allocation of scarce resources.[19] It is important to determine if programs and policies are reaching at-risk populations with the greatest needs.

To our knowledge, no published studies have examined how resources from various funding agencies are distributed to address childhood obesity among communities at risk. For example, are resources reaching all communities with need or only those with obesity prevalence above a selected threshold? What factors (such as the prevalence of childhood obesity and the percent of residents living in poverty) determine the allocation of scarce resources?

Our first objective is to determine whether resources allocated for early childhood obesity prevention programs have been directed toward neighborhoods with the greatest needs, specifically, those with a high prevalence of early childhood obesity, a high percentage of non-Hispanic Black and Hispanic residents, or those that are under-resourced socioeconomically. Our second objective is to identify specific neighborhoods that received fewer interventions than we would expect, based on their obesity prevalence and sociodemographic characteristics, to identify communities that policymakers should consider prioritizing. Our final objective is to determine the effect of the number of obesity-related interventions received by a neighborhood on its obesity prevalence among preschool-aged WIC children from 2005-2015 in LAC, considering two major events that may have affected obesity risk in this population, the Great Recession of 2008-2009 and the 2009 legislative change in the WIC food package.[178,179]

Methods

<u>Overview</u>

We created a database of interventions implemented in LAC from 2005 through 2015. Interventions were characterized using a typology developed to categorize childhood obesity prevention strategies,[150,154] and assigned to neighborhoods reached or targeted and years when the intervention was administered. As used in many studies, we used ZIP Codes as a proxy for neighborhoods.[111,180-183] For each ZIP Code, we summed the total number of intervention strategies implemented in a given year, providing a count of all intervention strategies (referred to as *intervention strategy count*) implemented in that ZIP Code per year. The intervention strategy count was linked to neighborhood-level prevalence of early childhood obesity and neighborhood-level sociodemographic characteristics.

<u>Study variables</u>

Intervention data were obtained from the ECOSyS Intervention Database (described in detail in CHAPTER 3). *Intervention strategy count* was conceptualized as the "amount" of intervention that a neighborhood (ZIP Code) received, and operationalized as the total number of intervention strategies implemented for each ZIP Code in a given year. An intervention strategy is defined as a plan of action that describes a method for achieving project objectives and producing defined outcomes. Intervention strategies used by each project were determined using a modified typology that identified 10 strategies, expanding on the 9 strategies identified by the original typology,[150,154] broadly categorized as macro-level or micro-level. A macro-level strategy is defined as one that does not directly target individuals but may affect the larger community such as a government policy or park facilities. A micro-level strategy is defined as one that directly targets individuals such as nutrition education.[150] A list of these strategies is provided in **CHAPTER 3, Table 3.1**. A program may have several projects and each project may use one or more (up to 10) intervention strategies. Each project that reaches or targets a ZIP Code contributes a minimum of 1 and a maximum of 10 intervention strategies to that ZIP Code in a given year. The intervention strategy count sums the total number of intervention strategies implemented in a ZIP Code in a given year. The hierarchical relationships among programs, projects, and intervention strategies, with definitions of these terms, can be seen in **CHAPTER 3, Figure 3.2**.

Obesity prevalence estimates were calculated from data provided by the LA County WIC Data Mining Project (described in detail in CHAPTER 3). *Neighborhood-level childhood obesity prevalence* was calculated as the percent of WIC-participating children ages 2 to 5 years who were obese in a ZIP Code during a 3-year period. We used a 3-year period to obtain stable estimates, similar to the ACS 3- and 5-year estimates of sociodemographic characteristics used by the Census Bureau[184] (e.g. 2003-2005 for the 2005 3-year estimate, 2004-2006 for the 2006 3-year estimate, etc.). Obesity status for children 2 to 5 years enrolled in WIC was determined using data on height, weight, and age. A child was categorized as obese if the child had a Body Mass Index (BMI) ≥95th percentile of CDC's sex-and age-specific growth reference values.[46] Neighborhood-level obesity prevalence was calculated by dividing the total number of unique children who were obese in a given ZIP Code for each 3-year period (e.g. between 2003 and 2005 for the 2005 3-year estimate) divided by the total number of unique WIC-enrolled children residing in that ZIP Code for that 3-year period. Each unique child had an obesity status for each year the child was enrolled in WIC, meaning during a 3-year period, each child could have up to 3 obesity status measures. For each 3-year period, a child had to be obese at least half of the time to be considered as obese. For example, if a child was obese in 2003, but not obese in 2004 or 2005, that child would be not be counted as obese for the 2003-2005 3-year period. Due to issues of confidentiality, of the 311 ZIP Codes in LAC, only ZIP Codes with at least 30 WIC children in each 3-year period were included in the study (n= 258). *Three-year neighborhood-level childhood obesity prevalence* was calculated for 2005-2015.

Sociodemographic data were obtained from the Census Bureau's Decennial Census and the American Community Surveys (ACS) at the level of ZIP Code Tabulation Areas (ZCTA), which are statistical geographic representations of ZIP Codes created by aggregations of census blocks.[157,159,185] *Neighborhood sociodemographic characteristics* were operationalized by three variables: (a) *neighborhood poverty* defined as the percent of persons living below the federal poverty line, (b) *neighborhood education* defined as the percent of residents 25 years or older without a high school degree, and (c) *neighborhood racial/ethnic minority composition* defined as the percent of residents who were non-Hispanic Black or Hispanic. Census data are available at the ZCTA level for 2000 and 2010 onwards. For this study, we used the 2000 Census, and the ACS 5-year estimates for 2010 through 2015. Sociodemographic data for each ZCTA were obtained in two ways. For the years 2010-2015, the data were obtained directly from the ACS 5-year estimates for each year. We used linear interpolation to estimate neighborhood sociodemographic values of interest for 2005-2009 using data from the 2000 Census and the 2010 ACS.

Statistical analyses

The minimum, maximum, and mean number of programs, projects, and intervention strategies were determined for each neighborhood (ZIP Code) for each year between 2005 and 2015. To visually examine the distribution of interventions, obesity, and sociodemographic characteristics across LAC, we mapped intervention strategy count, early childhood obesity prevalence, poverty levels, education levels, and racial/ethnic composition for each neighborhood in LAC for the years 2005, 2010, and 2015.

To determine whether resources allocated for obesity prevention programs have been directed toward neighborhoods with the greatest needs, we examined the association of *each* of the four neighborhood characteristics (childhood obesity, education, poverty, and minority composition) with intervention strategy count for the years 2005, 2010, and 2015. First, we ran four generalized linear models (GLMs) with a Gamma distribution and log link,[186] regressing intervention strategy count on each of the four neighborhood characteristics, separately, to examine the relationship of each of the four neighborhood characteristics with intervention strategy count. We then ran a full GLM model regressing intervention strategy count on neighborhood levels of early childhood obesity, poverty, education, and percent non-Hispanic Black or Hispanic to identify which neighborhood characteristics were most important for determining how resources were allocated for obesity prevention programs.

To identify specific neighborhoods that received fewer interventions than we would expect, we applied this full model to estimate each neighborhood's predicted intervention strategy count based on the neighborhood's obesity prevalence and sociodemographic characteristics. We then subtracted the observed intervention strategy count from the predicted intervention strategy count, and averaged this value over 2005 through 2015, for each neighborhood. We used this average difference (between predicted and observed intervention strategy count) to identify the top neighborhoods that received fewer interventions than expected based on their obesity prevalence and sociodemographic characteristics.

To determine the effects of the number of obesity-related interventions received by a neighborhood and its obesity prevalence, we ran fixed-effects linear models using ordinary least squares (OLS) to estimate the association between neighborhood-level early childhood obesity prevalence and intervention strategy count neighborhoods received in the preceding year (*time* $_{t-1}$). We conducted our analyses sequentially, first constructing an unadjusted model and then adjusting for the following time-variant variables: neighborhood levels of poverty, education, and minority composition (percent non-Hispanic Black or Hispanic), as well as prior-year obesity prevalence (*time* $_{t-1}$). We also included a dummy variable for the year 2008 and 2009 to control for the 2008-2009 Great Recession[187] and the 2009 WIC food package change.[188]

$$\begin{aligned} ObesityPrev_{it} &= \alpha_i + \beta_1 InterventionStrategyCount_{i(t-1)} + \beta_2 Poverty_{it} + \ \beta_3 Education_{it} \\ &+ \beta_4 Minority_{it} + \beta_5 Obesity_{i(t-1)} + \beta_6 Year2008Dummy_{it} \\ &+ \beta_7 Year2009Dummy_{it} + \mu_{it} \end{aligned}$$

Robust standard errors were calculated using the R package `lmtest`.[189] Sensitivity analyses were conducted to assess how robust our findings were after removing wealthy neighborhoods from all of our regression-based analyses. Findings from the sensitivity analyses were consistent with our findings from the main analyses. All statistical analyses were conducted using R version 4.0.3.[190] All maps were created using ArcGIS 10.5 (ESRI Redlands, CA).

Results

Interventions

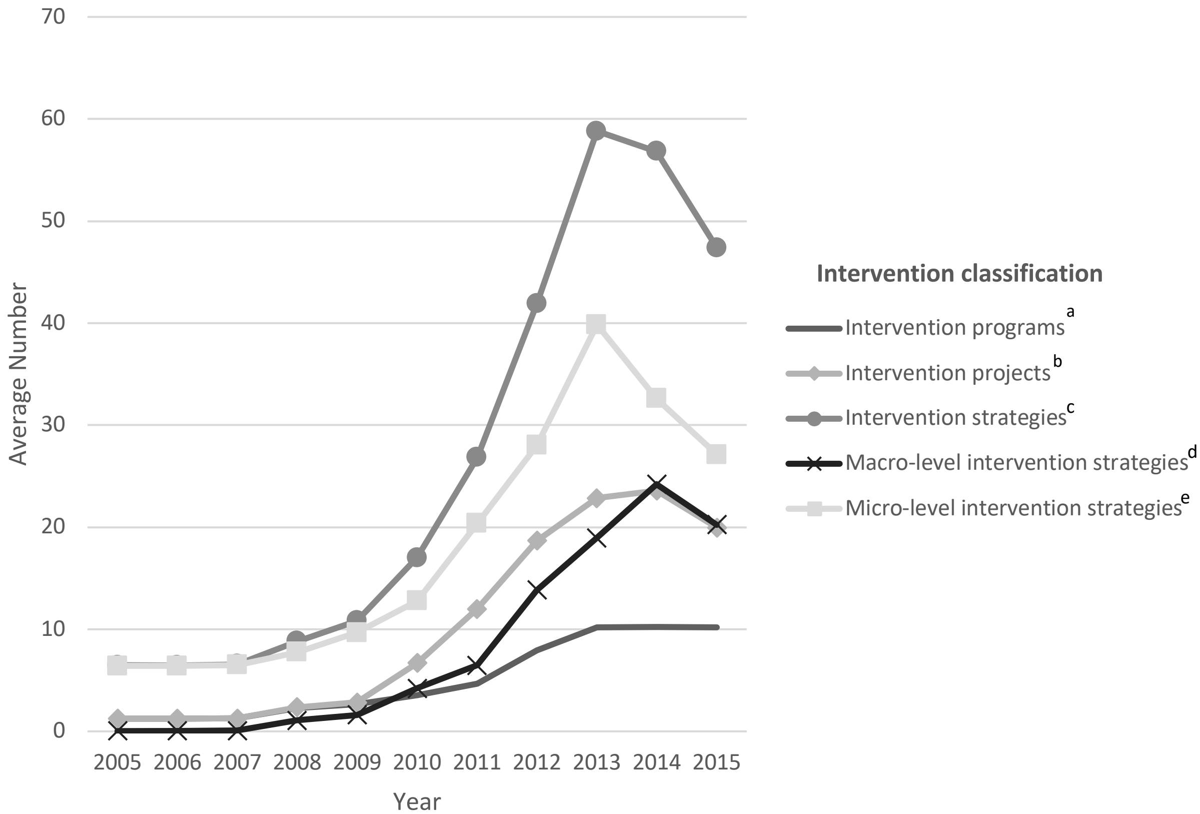
Table 4.1 and **Figure 4.2** provide a summary of interventions addressing obesity in LAC between 2005 and 2015. Between 2005 and 2007, 307 (of the 311 neighborhoods in LAC) received obesity-related interventions. By 2008, all neighborhoods received some form of obesity-related intervention. On average, each neighborhood had an intervention strategy count of about 6.5 between 2005 and 2008. The intervention strategy count per neighborhood steadily increased until 2013 when it peaked at 56.8, and then decreased to an average of 47.4 in 2015. Neighborhoods tended to receive more micro-level intervention strategies than macro-level intervention strategies.

Table 4.1. Summary of neighborhood-level interventions addressing obesity in Los Angeles County, 2005-2015

		2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Number of neighborhoods[a] with interventions		307	307	307	311	311	311	311	311	311	311	311
Number of intervention programs[b] per neighborhood	*Min*	0	0	0	2	2	2	3	6	8	8	8
	Max	5	5	6	7	8	10	11	14	19	20	19
	Mean	1.24	1.24	1.27	2.3	2.67	3.53	4.70	7.91	10.17	10.23	10.22
Number of intervention projects[c] per neighborhood	*Min*	0	0	0	2	2	3	8	15	20	20	16
	Max	5	5	6	7	10	23	32	40	42	42	35
	Mean	1.24	1.24	1.28	2.35	2.82	6.68	11.99	18.67	22.85	23.58	19.98
Number of intervention strategies[d] per neighborhood	*Min*	0	0	0	8	8	9	18	34	51	47	37
	Max	20	19	22	24	29	42	61	81	95	96	88
	Mean	6.46	6.45	6.57	8.82	10.82	17.01	26.85	41.88	58.81	56.84	47.37
Number of macro-level intervention strategies[e] per neighborhood	*Min*	0	0	0	1	1	2	4	12	17	22	18
	Max	3	2	3	4	6	16	23	31	34	37	30
	Mean	0.05	0.04	0.06	1.07	1.61	4.22	6.47	13.86	18.96	24.21	20.25
Number of micro-level intervention strategies[f] per neighborhood	*Min*	0	0	0	7	7	7	14	22	34	25	19
	Max	18	18	19	20	25	33	40	51	67	61	62
	Mean	6.42	6.42	6.51	7.75	9.66	12.79	20.38	28.02	39.85	32.62	27.12

[a]Neighborhood defined as ZIP Codes; Los Angeles County has 311 ZIP Codes
[b]A long-term managed portfolio of multiple projects to produce outcomes
[c]A project is short-term and designed to deliver a specified output within a specified time period and location, and may use various intervention strategies
[d]A plan of action that describes a method for achieving project objectives and producing defined outcomes
[e]Intervention strategies that indirectly affect the larger community and obesity-related behaviors and practices
[f]Intervention strategies that directly target individuals to modify obesity-related behaviors

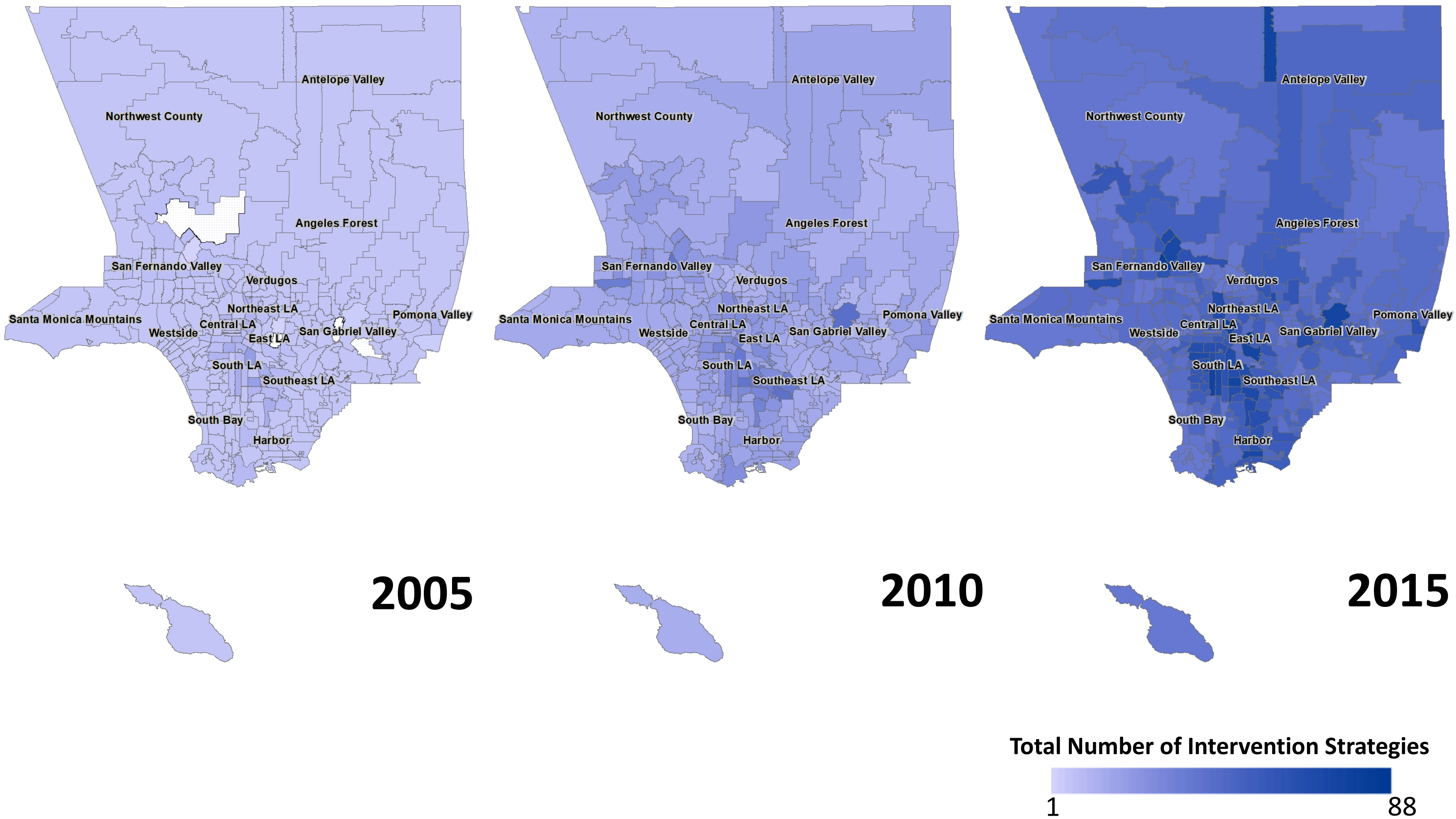
Figure 4.2. Average number of obesity-related interventions received by neighborhoods in Los Angeles County by intervention classification and year, 2005-2015



[a]A long-term managed portfolio of multiple projects to produce outcomes
[b]A project is short-term and designed to deliver a specified output within a specified time period and location, and may use various intervention strategies
[c]A plan of action that describes a method for achieving project objectives and producing defined outcomes
[d]Intervention strategies that indirectly affect the larger community and obesity-related behaviors and practices
[e]Intervention strategies that directly target individuals to modify obesity-related behaviors

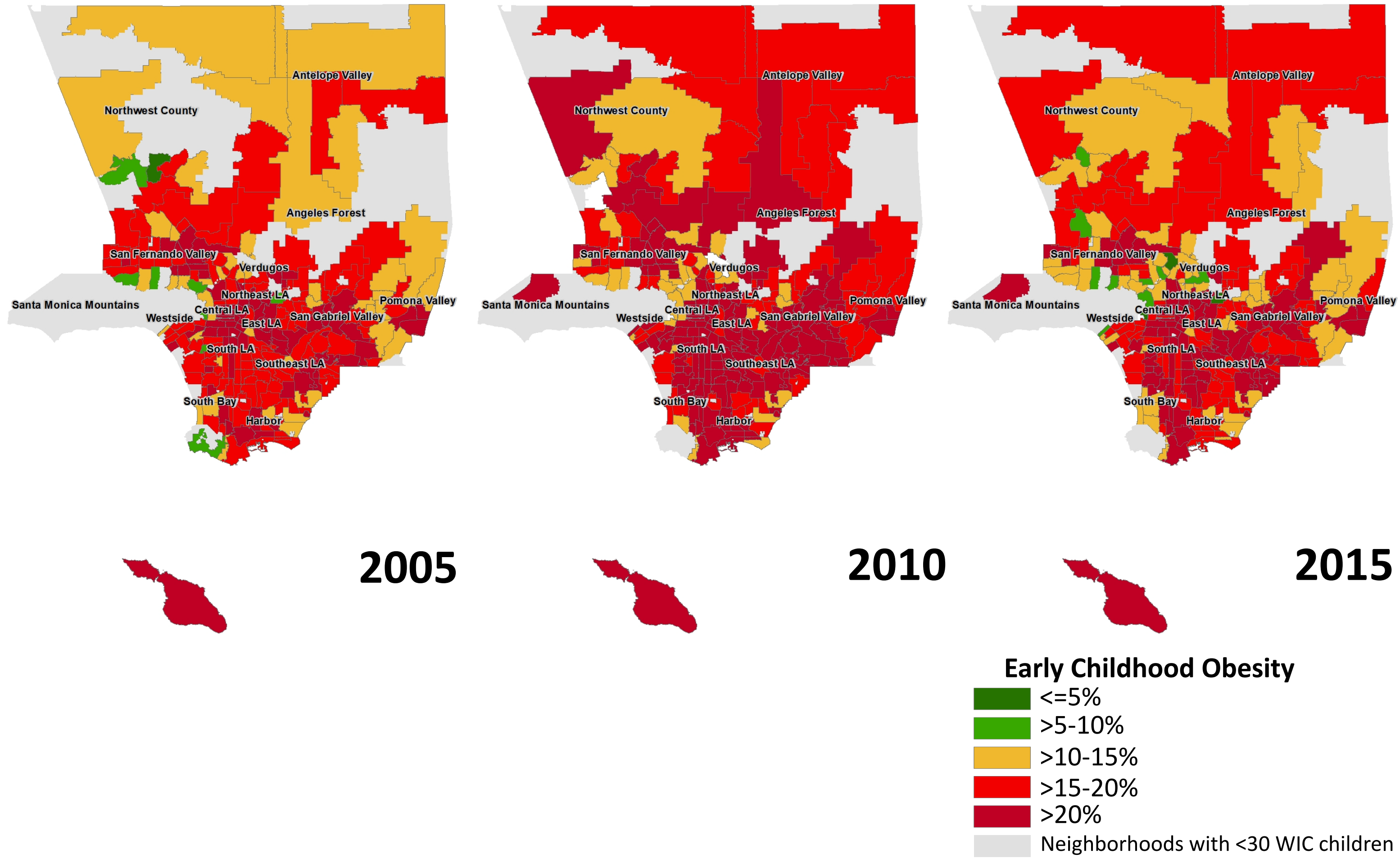
The distribution of intervention strategy counts across neighborhoods in LAC is shown in **Figure 4.3**. In 2005, the intervention strategy count tended to be higher in neighborhoods in South LA. Compared to 2005, intervention strategy count increased across all neighborhoods in LAC in 2010, and was highest among neighborhoods in South LA, Southeast LA, and parts of San Fernando and San Gabriel Valleys. By 2015, intervention strategy count continued to remain highest among neighborhoods in South LA, Southeast LA, and parts of San Fernando and San Gabriel Valleys, as well a few neighborhoods in Harbor, Antelope Valley, and Pomona Valley.

Figure 4.3. Distribution of intervention strategy count[a] across neighborhoods in Los Angeles County, 2005-2015



Neighborhood defined as ZIP Codes; Los Angeles County has 311 ZIP Codes
[a]Intervention strategy count operationalized as the total number of intervention strategies targeting obesity implemented for each ZIP Code in a given year; intervention strategy is defined as a plan of action that describes a method for achieving project objectives and producing defined outcomes

Figure 4.4. Prevalence of early childhood obesity[a] across neighborhoods in Los Angeles County, 2005-2015



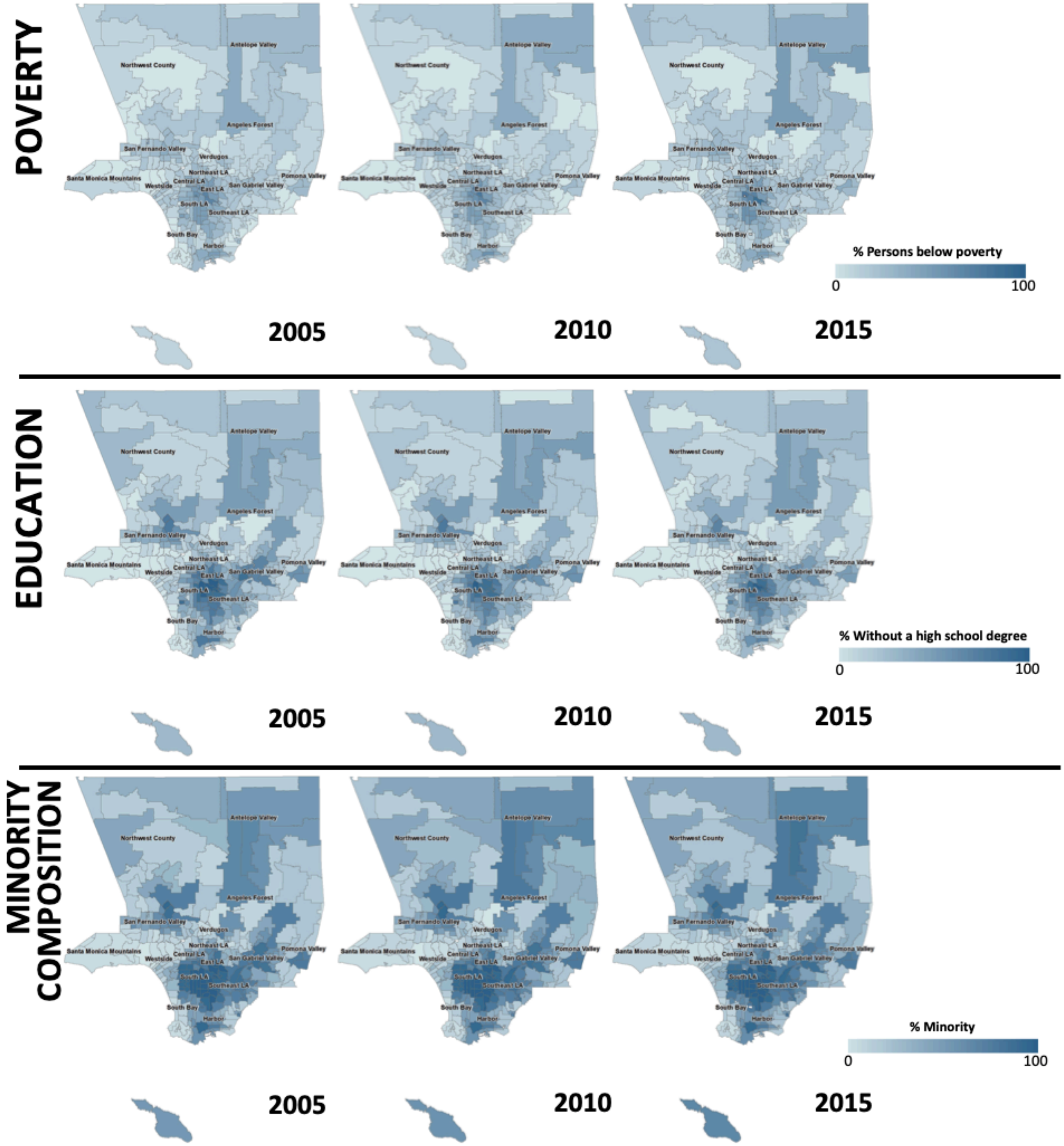
DATA SOURCE: PHFE WIC Data Mining Project
Neighborhood defined as ZIP Codes; Los Angeles County has 311 ZIP Codes
[a]Obesity prevalence averaged over 3 years; obesity status for children 2 to 4 years enrolled in the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) was defined as having a BMI ≥95th percentile of CDC's sex- and age-specific growth reference values; only ZIP Codes with at least 30 WIC children were included (n= 258)

Table 4.2. Summary of neighborhood-level early childhood obesity prevalence and sociodemographic characteristics across neighborhoods in Los Angeles County, 2005-2015 (n= 311 neighborhoods)

Neighborhood[a] characteristics		**Year**										
		2005	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**
Early childhood obesity (%)[b]	*Lowest*	5.00	2.44	6.67	7.81	3.33	5.71	5.41	7.32	6.56	5.71	4.43
	Highest	40.98	39.58	35.71	28.06	46.30	41.67	40.28	28.57	28.13	27.81	30.51
	Average	17.79	18.83	19.88	20.15	20.31	20.38	19.98	19.68	18.97	18.26	17.79
Persons below poverty (%)	*Lowest*	1.99	2.15	2.31	2.47	2.63	1.70	1.10	1.60	1.80	1.10	2.50
	Highest	58.05	62.32	66.59	70.87	75.14	54.10	68.90	78.80	72.40	64.10	65.20
	Average	15.24	15.33	15.42	15.51	15.57	13.96	14.33	15.30	15.77	16.30	16.24
Less than high school degree (%)	*Lowest*	0.00	0.00	0.00	0.43	1.61	1.20	0.00	0.00	0.00	0.00	0.00
	Highest	70.17	69.41	69.05	69.02	69.00	64.60	69.60	71.30	72.80	74.20	69.80
	Average	23.15	22.68	22.22	21.75	21.29	20.21	20.02	19.89	19.57	19.31	18.83
Minority (%)[c]	*Lowest*	4.47	4.69	4.91	4.81	4.11	4.20	3.20	4.00	2.80	1.70	0.70
	Highest	98.64	98.62	98.64	98.66	98.68	98.60	98.90	98.80	98.70	98.90	98.90
	Average	44.94	45.09	45.25	45.40	45.56	46.12	46.06	46.17	46.33	46.48	46.62

DATA SOURCES: PHFE WIC Data Mining Project; 2000 Census, 2011-2015 American Community Surveys (U.S. Census Bureau)
[a]Neighborhood defined as ZIP Codes for early childhood obesity data and ZIP Code Tabulation Areas (ZCTAs) for sociodemographic data
[b]3 year estimates; obesity status for children 2 to 4 years enrolled in the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) was defined as having a BMI ≥95th percentile of CDC's sex- and age-specific growth reference values; only ZIP Codes with at least 30 WIC children were included (n= 258)
[c]Minority defined as Non-Hispanic Black/African American or Hispanic/Latino
NOTE: Los Angeles County has 311 neighborhoods (ZIP Codes/ZCTAs)

Figure 4.5. Distribution of sociodemographic characteristics across neighborhoods in Los Angeles County, 2005-2015



DATA SOURCES: 2000 Census, 2010 and 2015 American Community Surveys (U.S. Census Bureau)
Neighborhood defined as ZIP Codes; Los Angeles County has 311 ZIP Codes
[a]Minority defined as Non-Hispanic Black/African American or Hispanic/Latino

Early childhood obesity prevalence

The mean prevalence of early childhood obesity per neighborhood steadily increased from 17.8% in 2005 to 20.4% in 2010, then slowly decreased back to 17.8% in 2015. The distributions of early childhood obesity prevalence across neighborhoods in LAC are shown in **Figure 4.4**. In 2005, obesity prevalence was highest among neighborhoods in Central and East LA, and parts of Southeast LA, Harbor, as well as San Fernando and San Gabriel Valleys. Overall, obesity prevalence increased throughout LAC in 2010 with obesity prevalence greater than 20% in most neighborhoods. In 2015, though obesity prevalence decreased throughout LAC, obesity prevalence remained above 20% among neighborhoods in Central LA, East LA, South LA, Southeast LA, San Gabriel Valley, and parts of Harbor, San Fernando Valley, and Pomona Valley.

Neighborhood sociodemographic characteristics

Neighborhood levels of poverty, education, and minority composition slightly increased between 2005 and 2015. On average, neighborhoods in LAC had about 15% of its residents living below poverty, about 20% of residents 25 years or older without a high school degree, and about 45% of residents were non-Hispanic Black or Hispanic **(Table 4.2).** The distributions of sociodemographic characteristics across neighborhoods in LAC are provided in **Figure 4.5**. Poverty, low levels of education, and high percentages of non-Hispanic Black or Hispanic residents were largely concentrated in neighborhoods in South LA, East LA, and Southeast LA, and parts of Harbor in 2005, 2010, and 2015.

Characteristics of neighborhoods where interventions were implemented

The associations of neighborhood early childhood obesity prevalence and neighborhood sociodemographic characteristics with intervention strategy count are shown in **Table 4.3**. In the unadjusted models, higher intervention strategy count was significantly associated with each neighborhood characteristic evaluated. Higher obesity prevalence was positively associated with intervention strategy count. In 2005, for each 1% increase in obesity prevalence, the mean intervention strategy count would be expected to increase by a factor of 1.012 (95% confidence interval (CI): 1.001, 1.022). Neighborhoods with higher levels of poverty were likelier to receive more intervention strategies (higher intervention strategy count). For each 1% increase in the percent of persons living below poverty, the mean intervention strategy count would increase by a factor of 1.009 (95% CI: 1.006, 1.012) in 2005. Intervention strategy count was negatively associated with education levels, such that as education levels decreased, intervention strategy count increased. For each 1% increase in the percent of residents without a high school degree, the mean number of intervention strategies a neighborhood received would increase by a factor of 1.005 (95% CI: 1.003, 1.007) in 2005. Finally, higher intervention strategy count was associated with higher percent of minority residents in a neighborhood. For each 1% increase in the percent of minority residents, the mean intervention strategy count increased by a factor of 1.003 (95% CI: 1.002, 1.004) in 2005. The significance of each of these relationships held in 2010 and 2015. In the full model that included all four neighborhood-level variables (obesity prevalence and neighborhood sociodemographic characteristics), intervention strategy count was significantly associated with neighborhood-level poverty and minority composition. Neighborhoods with higher poverty levels and higher percentages of non-Hispanic Black and

Hispanic residents received more intervention strategies. For each 1% increase in the percent of residents living below poverty, the mean intervention strategy count increased by a factor of 1.011 (95% CI: 1.007, 1.018), 1.007 (95% CI: 1.001, 1.014), and 1.006 (95% CI: 1.003, 1.010) in 2005, 2010, and 2015, respectively, while holding all other variables in the model constant. For each 1% increase in the percent of minority residents in a neighborhood, the mean intervention strategy count increased by a factor of 1.003 (95% CI: 1.001, 1.006) and 1.004 (95% CI: 1.003, 1.005) in 2010 and 2015, respectively, while holding all other variables in the model constant.

Table 4.3. Generalized linear models (GLMs) with a Gamma distribution and log link: Association of neighborhood early childhood obesity prevalence and neighborhood sociodemographic characteristics with intervention strategy count[a]

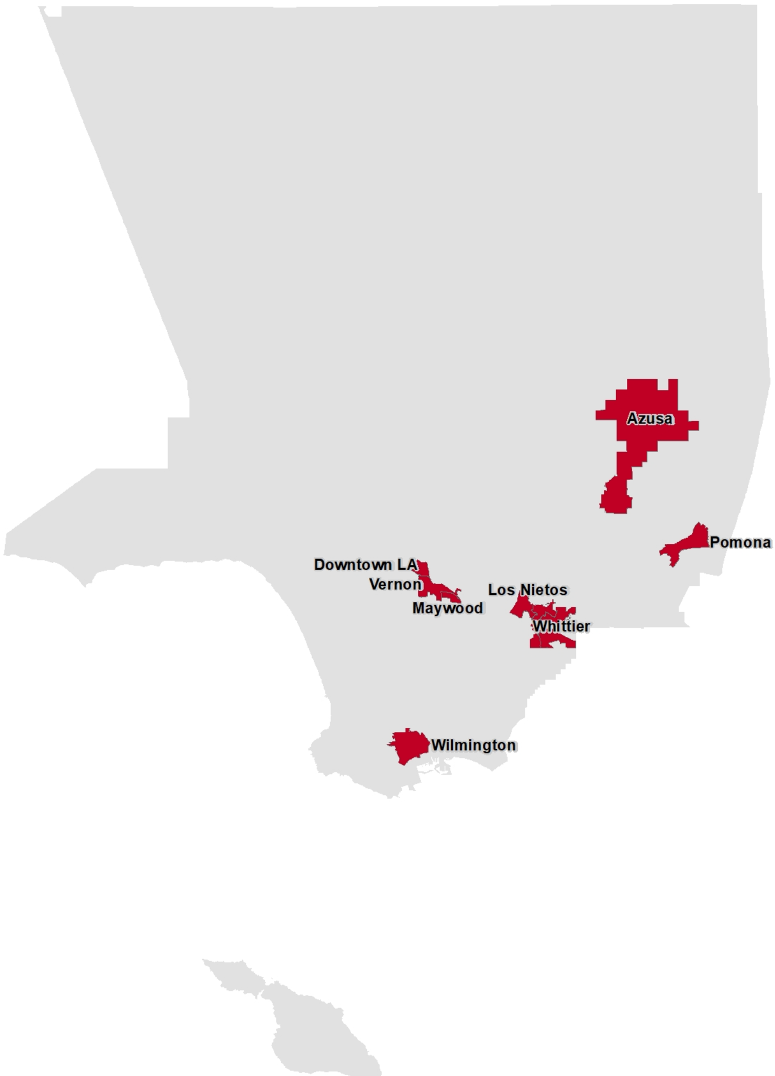
		Gamma GLM model (log link), unadjusted	Gamma GLM model (log link), full model[b]
		Mean Ratios exp(β) (95% CI)	Mean Ratios exp(β) (95% CI)
% Early childhood obesity[c]	2005	**1.012 (1.001, 1.022)**	0.997 (0.986, 1.008)
	2010	**1.025 (1.016, 1.034)**	0.994 (0.984, 1.005)
	2015	**1.017 (1.012, 1.022)**	0.997 (0.991, 1.003)
% Persons below poverty	2005	**1.009 (1.006, 1.012)**	**1.011 (1.005, 1.018)**
	2010	**1.017 (1.013, 1.021)**	**1.007 (1.001, 1.014)**
	2015	**1.009 (1.007, 1.011)**	**1.006 (1.003, 1.010)**
% Less than a high school degree	2005	**1.005 (1.003, 1.007)**	0.998 (0.993, 1.004)
	2010	**1.011 (1.009, 1.013)**	1.003 (0.998, 1.009)
	2015	**1.008 (1.007, 1.009)**	0.999 (0.996, 1.002)
% Minority[d]	2005	**1.003 (1.002, 1.004)**	1.001 (0.998, 1.004)
	2010	**1.006 (1.005, 1.007)**	**1.003 (1.001, 1.006)**
	2015	**1.005 (1.004, 1.005)**	**1.004 (1.003, 1.005)**

CI, confidence interval. Statistically significant values are in bold. Neighborhood defined as ZIP Codes.
[a]Intervention strategy count operationalized as the total number of intervention strategies targeting obesity implemented for each ZIP Code in a given year; intervention strategy is defined as a plan of action that describes a method for achieving project objectives and producing defined outcomes
[b]$InterventionStratCount = \beta_0 + \beta_1 ObesityPrevalence + \beta_2 Poverty + \beta_3 Education + \beta_4 Minority$
[c]Obesity prevalence averaged over 3 years; obesity status for children 2 to 4 years enrolled in the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) was defined as having a BMI ≥95th percentile of CDC's sex- and age-specific growth reference values; only ZIP Codes with at least 30 WIC children were included (n= 258)
[d]Minority defined as Non-Hispanic Black/African American or Hispanic/Latino

Neighborhoods that received fewer interventions than expected

A map of the neighborhoods that received fewer interventions than would be expected, based on neighborhood-level obesity prevalence and sociodemographic characteristics, is provided in **Figure 4.6**. We identified a cluster of neighborhoods in and around Downtown/Southeast LA (Vernon, Downtown LA, Maywood) and San Gabriel Valley (Los Nietos and Whittier), as well as the neighborhoods of Wilmington, Pomona, and Azusa. The identified neighborhoods are lower income neighborhoods with lower education levels compared to the average neighborhood in LAC. All of these neighborhoods have high Hispanic populations, with the exception of Downtown LA that is highly diverse with about 35% Hispanic, 20% Black, and 20% Asian residents.[174]

Figure 4.6. Neighborhoods that received a lower intervention strategy count than expected based on neighborhood levels of obesity and sociodemographic characteristics



Neighborhoods were identified by comparing neighborhoods' observed intervention strategy count with its predicted intervention strategy count based on its level of obesity prevalence and sociodemographic characteristics and the model:

$$InterventionStrategyCount = \beta_0 + \beta_1 ObesityPrevalence + \beta_2 Poverty + \beta_3 Education + \beta_4 Minority$$

Effect of the number of obesity-related interventions neighborhoods received on obesity prevalence

The results of our fixed-effects model to evaluate the association between the number of obesity-related interventions neighborhoods received (intervention strategy count) and neighborhood obesity prevalence in the following year are presented in **Table 4.4**. Higher intervention strategy count was associated with lower neighborhood-level early childhood obesity prevalence in the following year. In the unadjusted model, each additional intervention strategy a neighborhood received was significantly associated with lower early childhood obesity prevalence in the following year (β= -0.014; 95% CI: -0.021, -0.001). After adjusting for neighborhood sociodemographic characteristics, prior-year obesity prevalence, the 2008-2009 Great Recession, and the 2009 WIC food package change, this relationship remained with higher intervention strategy count associated with lower neighborhood-level early childhood obesity prevalence. Adjusted for time-variant variables, each additional intervention strategy a neighborhood received was significantly associated with lower early childhood obesity prevalence in the following year (β= -0.023; 95% CI: -0.031, -0.016).

Table 4.4. Fixed-effects models of the association between early childhood obesity prevalence and intervention strategy count[a]

	Fixed-effects model, unadjusted	Fixed-effects model, adjusted[b]
	β (95% CI)	β (95% CI)
Intervention strategy count$_{t-1}$	**-0.031 (-0.038, -0.025)**	**-0.030 (-0.036, -0.024)**

CI, confidence interval. Statistically significant values are in bold. Neighborhood defined as ZIP Codes.

[a]Intervention strategy count operationalized as the total number of intervention strategies targeting obesity implemented for each ZIP Code in a given year; intervention strategy is defined as a plan of action that describes a method for achieving project objectives and producing defined outcomes

[b]$ObesityPrev_{it} = \alpha_i + \beta_1 InterventionDose_{i(t-1)} + \beta_2 Poverty_{it} + \beta_3 Education_{it} + \beta_4 Minority_{it} + \beta_5 Obesity_{i(t-1)} + \beta_6 Year2008Dummy_{it} + \beta_7 Year2009Dummy_{it} + \mu_{it}$

Discussion

As a response to the obesity epidemic, place-based interventions to address early childhood obesity have been implemented by various organizations across LAC. Our study found (1) increasing interventions levels over the period of study (2005-2016), (2) neighborhoods with the highest levels of poverty and a higher proportion of non-Hispanic Black or Hispanic residents generally received more resources, (3) a small group of neighborhoods that received fewer interventions that expected, and (4) neighborhoods that received more intervention saw significantly lower prevalence of early childhood obesity in the following year.

We found that intervention strategy count has steadily increased across neighborhoods in LAC since 2005. In LAC, neighborhoods with higher levels of poverty and higher percentages of non-Hispanic Black and Hispanic residents generally received more interventions. These findings track with large-scale initiatives to address early childhood obesity that have been implemented in LAC. For example, in 2010, First 5 LA (F5LA)— one of 58 county commissions in California created by Proposition 10, identified 14 Best Start Communities (BSCs) throughout LAC that are racially and ethnically diverse, face critical issues, such as poverty and low school performances, but also have a strong network of local leaders and organizations dedicated to their communities.[175] Since 2009 and 2010, F5LA has been heavily investing in BSCs to support community partnerships that help to promote positive outcomes for young children and their families, including reductions in early childhood obesity.[175] Through these partnerships, F5LA has invested in programs using multiple interwoven intervention strategies to strengthen both the capacity of families to raise healthy children and the capacity of communities and broader systems to support healthy families. An example of such a program is The Welcome Baby

Program, which involved the implementation of six strategies at the macro and micro levels including Baby-Friendly Hospital policies, home visitations, staff training, breastfeeding education, as well as counseling and referrals. Breastfeeding has been observed to reduce risk of childhood overweight and obesity.[191]

While we found that resources from the major initiatives generally targeted neighborhoods with the greatest needs, we also identified a small group of neighborhoods that received fewer interventions from the major initiatives we evaluated than expected. About half of these neighborhoods had >85% minority populations, high poverty rates, and low educational levels.[52] Early childhood obesity prevalences in these neighborhoods were high, averaging about 20%. However, many neighborhoods in LAC had higher prevalences of early childhood obesity, underscoring the importance of coordinated efforts to distribute limited resources efficiently and effectively.

Neighborhoods that received higher intervention strategy count saw a reduction in early childhood obesity prevalence in the following year. For example, in 2015, neighborhoods in LAC received an average of 50 intervention strategies. Based on our study findings, this would translate to a reduction in obesity prevalence of about 1% points in the following year.

Several different approaches to population health interventions have been proposed over the years. Lalonde's "high-risk" approach aims to target individuals at high-risk for a disease in order to maximize health.[124] Rose's "population" approach aims to target the entire population regardless of risk factor, disease status, or need. While this has been shown to be more effective in reducing disease risk at the population level,[126] Frohlich and Potvin contend that disparities in health may be exacerbated by such interventions, and propose the "vulnerable population" approach to target disadvantaged populations, rather than the entire population, to reduce health disparities.[130]

Our findings suggest that community-based prevention approaches that target under-resourced communities, rather than high-risk individuals or entire populations, and designed to modify both individual behaviors and the environmental contexts in which they develop, have the potential to reduce rates of early childhood obesity— supporting Frohlich and Potvin's approach.[130] Since children from low-income families are more likely to be obese than children from high income families,[49] interventions that help reduce rates of early childhood obesity in low-income communities can help reduce overall rates of obesity among preschool-aged children in the US. Our findings support the importance of place-based interventions in reducing rates of obesity. Though neighborhoods that received more intervention strategies saw greater declines in obesity prevalence, in order to allocate resources more efficiently we still need to understand if and when these effects plateau, understand what combination of intervention strategies are most effective in reducing population-level rates of obesity, and conduct cost-effectiveness studies.

Study strengths and limitations

Creating an intervention database that classified programs and projects facilitated the concept of "intervention strategy count" which allowed us to quantify exposure of a community to various intervention strategies implemented simultaneously by obesity-related interventions. Our findings of the negative relationship between intervention strategy count and subsequent neighborhood obesity prevalence contributes to the mixed evidence on the impact of place-based interventions on obesity at the population-level.[21-23]

Our study has several limitations. First, our intervention database is not comprehensive of all obesity-related interventions that took place in LAC. It focused on interventions implemented by major funders and health organizations tackling obesity in LAC, and WIC clinics in 8 regions of LAC where the majority of WIC families reside (see **CHAPTER 3, Figure 3.1**). Second, while the intervention strategy count allows for the quantification of exposure to various strategies, it does not consider the reach of a strategy. Third, intervention data were available only at the ZIP Code level, necessitating neighborhoods to be defined by ZIP Codes; these are relatively large geographic spaces and consequentially more likely to display heterogeneous neighborhood effects.[192]

Conclusion

As considerable amounts of funding are being spent to tackle obesity, it is important to determine if programs and policies are reaching at-risk populations with the greatest needs. In LAC, we found that obesity-related interventions have been targeting under-resourced communities. Specifically, poorer neighborhoods and neighborhoods with a high percentage of non-Hispanic Black or Hispanic residents have received more resources for obesity prevention. In a community setting, where neighborhoods receive a variety of simultaneous interventions, it is challenging to evaluate the effectiveness of place-based obesity prevention efforts. This work demonstrates the usefulness of an intervention database, such as the ECOSyS Intervention Database which contains data on all major initiatives addressing obesity at the community-level in LAC since 2003. Using intervention strategy count as a measure of the “amount” of intervention a neighborhood was exposed to, we found that neighborhoods that received more intervention strategies saw greater declines in obesity prevalence. Interventions that help reduce rates of early childhood obesity among low-income children can play a key role in helping to reduce overall rates of obesity among preschool-aged children in the US. Future research should evaluate what types of strategies and specific interventions are effective in reducing obesity at the neighborhood level.

CHAPTER 5: A practical application of machine learning techniques to identify neighborhoods most burdened by early childhood obesity

Introduction

Obesity is a serious problem in the United States (US), costing an estimated $147 billion (in 2008 dollars) in obesity-related medical care costs every year.[193] Consequently, obesity prevention is a major public health priority, particularly among children. Once childhood obesity is established, it is difficult to reverse through interventions,[12,13] and tracks into adulthood.[7,194] Childhood obesity starts very early in life,[3] and disproportionately affects children from low-income families.[49] In developed countries such as the US, children from lower socioeconomic status (SES) households are at increased risk of being obese compared with their higher SES counterparts.[195] This socioeconomic patterning of obesity has been observed across a number of socioeconomic markers, including parental education, family income, parental occupation, and the socioeconomic characteristics of area of residence.[164,196] In the US, there are also racial/ethnic disparities in childhood obesity with obesity more prevalent among Hispanic (25.8%) and non-Hispanic Black (22%) children compared with their white (14.1%) and Asian (11%) peers.[3] Disparities in childhood obesity based on SES as well as race/ethnicity are likely to further contribute to health disparities in adulthood in the US.[197,198]

Biologically, obesity is due to energy intake that exceeds energy expenditure, which is largely determined by diet and levels of physical activity (PA).[9] Diet and PA are the modifiable behaviors that have been traditionally targeted by many health education and promotion programs.[11,199] However, in the early 1990s, evidence on the influences of the built and social environment on these two behaviors suggested that access to healthy food and opportunities for

PA were limited in under-resourced neighborhoods.[123,128] A body of literature has since accumulated to support the need for place-based interventions. Evidence on the effectiveness of such interventions, while increasing, is still limited, but have shown some promise for preventing obesity. [21-24]

According to the socioecological model,[14] individual-level behaviors are contextualized by the individual's environment. The neighborhood built and social environment provide the context that can enable healthy behaviors or hinder them, and a neighborhood that promotes obesity is referred to as obesogenic.[17,81,82]

Broadly defined, the built environment encompasses human-made or human-modified aspects of a person's surroundings. The neighborhood built environment can affect energy balance by presenting opportunities or barriers to access healthy foods and opportunities or barriers for PA. Individuals living in areas that lack access to grocery stores with fresh fruits and vegetables, are more heavily targeted for advertising of unhealthy foods, and have easier access to fast-food restaurants may be more likely to adopt unhealthier diets.[82,84-87] Similarly, neighborhood residents' accessibility to recreational facilities such as parks and playgrounds, as well as the "walkability" of a neighborhood, that is, a neighborhood that provides a range of local facilities within walking or cycling distance and has supportive infrastructure like well-maintained sidewalks, can shape residents' levels of PA.[89-92]

The neighborhood social environment may influence food consumption and PA through neighborhood social capital, collective efficacy, and social norms. Neighborhood social capital—the quality and quantity of social resources in a community,[97] and collective efficacy— the mutual trust between neighborhood residents that enable them to collectively work together for the good of the community,[98,99] strongly predict neighborhood crime.[98,100] Increased crime may discourage residents from spending time outdoors and constrain opportunities for active living. Social norms, defined as explicit or implicit rules that guide, regulate, proscribe, and prescribe social behavior in particular contexts,[101] may influence behaviors related to eating and PA.[82,102,103] Neighborhood norms are, in part, a consequence of the sociodemographic characteristics of the neighborhood's residents. The greater the concentration of like-minded people, the stronger the norms, and the greater the exposure of residents to these norms.[104] For example, portion sizes offered by food outlets may be influenced by local norms that dictate a "normal" portion size. How "out of place" residents feel exercising in public areas, such as doing yoga in a park, may also be influenced by local norms. The neighborhood social environment can affect obesity by either supporting or discouraging obesity-related behaviors.

The neighborhood sociodemographic composition largely shapes the built and social environment of neighborhoods. It has consistently been demonstrated that living in under-resourced neighborhoods increases obesity risk, and these relationships hold using a broad range of neighborhood socioeconomic indicators including unemployment rates, area income and education, percent in poverty, and different indices of "community disadvantage" or "neighborhood socioeconomic status".[134,136,141,200] In the US, there is a strong correlation between neighborhoods that are under-resourced and minority neighborhoods due to the role that residential segregation has played in maintaining differences in SES by race.[201] Minority neighborhoods refer to neighborhoods with predominantly Black or Hispanic residents. Neighborhoods with more low SES residents and Black or Hispanic residents are likelier to have fewer and poorer built and social resources compared to neighborhoods with more higher SES residents or white residents.[134,136,141,200] The lack of supportive resources can restrict access to healthy food and opportunities for PA, while also negatively impacting social norms around these behaviors.[91] The availability and types of businesses, including food outlets and for-profit recreation facilities are often related to neighborhood sociodemographic factors.[105-107] In the US, lower income and minority neighborhoods have fewer supermarkets, fruit and vegetable markets, parks, sports facilities, and bike paths, and greater access to fast food restaurants.[107-111] Furthermore, higher prices for fresh produce have been found in areas of concentrated poverty compared to more affluent areas.[202] Under-resourced neighborhoods are also likelier to suffer from higher crimes rates.[98,113,114] According to social disorganization theory, neighborhood characteristics such as poverty, ethnic heterogeneity, high residential turnover rates, low homeownership rates, and concentration of recent immigrants may make it more difficult for residents to establish social ties, and build social capital and collective efficacy.[104,112] It has been shown that lack of social capital and collective efficacy strongly predict increased rates of neighborhood crime.[98,100]

A body of literature supports the relationship between neighborhood-level socioeconomic and sociodemographic characteristics with obesity.[107,134-136,141] Based on studies conducted in the US, neighborhood-level sociodemographic variables characteristics important for obesity include: racial/ethnic composition, education levels, marital status, income levels, poverty status, public assistance participation, employment status, place of birth, housing occupancy, age of homes in the neighborhood, and vehicle ownership.[107,134-137,141,165,180,203-237]

Identifying communities most burdened by obesity

In CHAPTER 4, we showed disparities in obesity prevalence across neighborhoods in Los Angeles County (LAC). Neighborhood-level prevalence of early childhood obesity among low-income children in LAC ranged from a low of 4.43% to a high of 30.51% in 2015. To reduce health disparities, Frohlich and Potvin[130] advocate for a "vulnerable population" approach whereby interventions should focus on populations that need it the most. To reduce disparities in obesity prevalence, communities that have the highest prevalences of early childhood obesity should be targeted. Identifying communities most burdened by obesity also allows obesity control programs to be tailored to address the specific needs of those communities and leverage on communities' existing resources. Such programs are more likely to be effective compared to broadly based programs targeting entire populations, such as programs delivered at the entire county or state.[131-133] However, current data limitations may thwart the efforts of funders and policymakers to identify these communities.

Identification of communities with the greatest needs for obesity prevention efforts require timely data on childhood obesity prevalence at the appropriate level of geography. We currently rely on surveillance data that are collected by national and state health surveys or surveillance systems for obesity prevalence estimates. Among publicly available data sources, obesity data for young children are available from the National Health and Nutrition Examination Survey (NHANES), the WIC Participant and Program Characteristics (WIC PC) survey, the California Health Interview Survey (CHIS), the North Carolina Child Health Assessment and Monitoring Program (NC-CHAMP), the Panel Study of Income Dynamics (PSID), and the Panel Study of Income Dynamics Child Development Supplement

(PSID/CDS).[55,56] These data sources, however, have significant lag times between data collection and data availability, and often have gaps between waves of data collection. Furthermore, obesity data from these sources are only available at the national or state level, with the exception of CHIS that provides estimates at the county level.[140] While these surveys and surveillance systems provide important information about the prevalence and trends of childhood obesity in the US at the national and state levels, the need to strengthen local public health data systems for surveillance at the neighborhood or small-area levels has been widely recognized.[238]

Local public health authorities and advocates are often in need of information about their own communities in order to appropriately prioritize, develop, and deliver appropriate interventions. Consequently, many cities, towns, and regions in the US have their own surveillance data available at smaller levels of geography to better monitor, evaluate, and improve the health of the populations they directly serve. For example, in LAC, data on obesity prevalence for WIC-enrolled preschool-aged children are available at the county level, as well as for smaller geographic units including census tracts and ZIP Codes.[54] The LA County WIC Data Mining Project provides access to a unique administrative database that includes data on child's age, ZIP Code of residence, height and weight for the duration participants' receive WIC services (from the prenatal period through the child's fifth birthday) for every year since 2003.[54] LAC is the only county in the US that is able to electronically aggregate and analyze WIC data across all WIC agencies in the county.[54] However, as with most surveillance systems, the lag times between data collection and data availability continue to prevent funders and policymakers from accessing quick, up-to-date estimates of obesity prevalence across communities. The predictive capabilities of machine learning may provide an opportunity to determine if existing data can be used to quickly identify communities in need of obesity control and overcome the lag times in data availability. Access to data from the LA County WIC Data Mining Project provides a unique opportunity to examine this research question.

Machine learning and public health

In recent years, the use of machine learning in public health research has grown exponentially,[239] and has been applied to research in areas spanning from disease diagnosis[240] to predictions of mortality risk,[241] air pollution,[242] overweight and obesity risk,[243-246] and state-level prevalence of non-communicable diseases.[247] Machine learning models are often able to provide more valid and accurate predictions than traditional approaches.[247,248] Health outcome estimates, such as obesity prevalence estimates, that are predicted from machine learning models, have the potential to play an important role in strategic decision making if they are able to achieve sufficient precision.[249] Predicted health outcome estimates, based on readily available data, may be able to address the issue of lag times between data collection and data availability observed with public health surveillance systems.

In the US, sociodemographic data are routinely collected by the US Census Bureau, and the data are easily accessible to the public at various levels of geography. As discussed previously, neighborhood-level sociodemographic factors are determinants of obesity.[107,134-136,141] Social, economic, and demographic characteristics of the environment can be quantified using spatial sociodemographic data provided by the US Census and the American Community Surveys (ACS). Unsurprisingly, there is a lag time between data collection and data availability, which is about two years, depending on the level of geography.[250]

The rationale for this study is to estimate future prevalence of early childhood obesity for neighborhoods in LAC using existing data. These estimates can provide up-to-date obesity prevalence estimates, which can then be used for targeted public health programs and interventions. LAC is the most populous county in the US with an ethnically diverse population of over 10 million residents.[52] About half of all infants and children under 5 years, or over half a million children, are enrolled in WIC.[53] Obesity prevalence among WIC-enrolled children 2 to 5 years in LAC has consistently been higher than that of the overall US population of preschool-aged children and overall WIC-enrolled population of children 2 to 5 years in the US, reaching a peak of about 20% in 2008, and then decreasing to 18% in 2016.[54]

Rapid identification of communities with the highest burden of obesity can help funders and policymakers allocate resources and prioritize these communities for intervention, thereby working towards reducing disparities in obesity. Using machine learning, Luo et al. used a small dataset of state-level sociodemographic characteristics from the census to model and reasonably predict state-level prevalence estimates for chronic diseases, including obesity.[247] To overcome the lag time between data collection and data availability of up-to-date obesity prevalence estimates, the overall aim of this study is to accurately predict future prevalence of early childhood obesity at the neighborhood level based on readily available data. We hypothesized that machine learning could be an effective approach to predicting future prevalence of early childhood obesity using the most up-to-date obesity prevalence estimates and sociodemographic data available. We compared the predictive accuracy of: (1) a machine learning model to predict future prevalence of early childhood obesity using both the latest obesity prevalence estimates and sociodemographic data available, (2) a machine learning model to predict future prevalence of early childhood obesity using only the latest sociodemographic data available, and (3) using only the latest obesity prevalence estimate as a proxy for future prevalence of early childhood obesity. We hypothesize that using both the latest obesity prevalence estimates and sociodemographic data available would most accurately predict future prevalence of early childhood obesity at the neighborhood level than either alone.

Methods

Overview

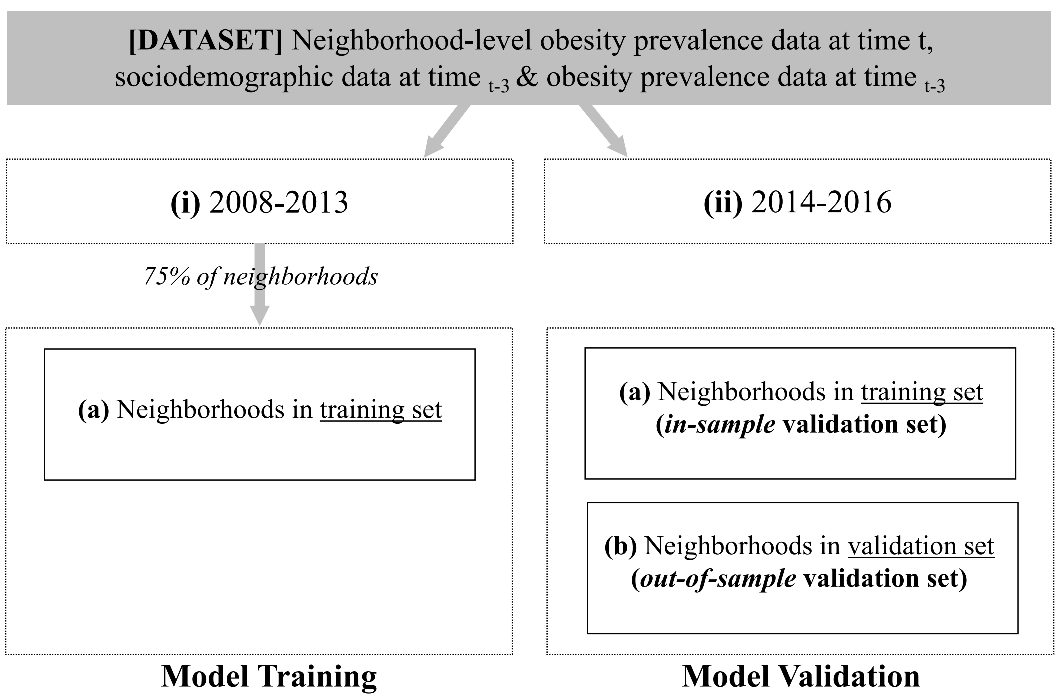
To predict future prevalence of early childhood obesity at the neighborhood level, we linked obesity prevalence estimates at *time t* with obesity prevalence estimates from the preceding 3 years (*time* $_{t-3}$), and sociodemographic data from the preceding 3 years (*time* $_{t-3}$). We used data from the preceding 3 years to represent the most recent obesity prevalence estimates and sociodemographic data available, which accounts for a comfortable lag time in data availability. For example, obesity prevalence in 2008 was linked with 2005 estimates of obesity prevalence and 2005 sociodemographic data, obesity prevalence in 2009 was linked with 2006 estimates of obesity prevalence and 2006 sociodemographic data, etc.

Machine learning involves training and validating a model. Model "training" refers to providing a machine learning algorithm with data to help the algorithm identify and "learn" the best values for all attributes involved. For example, for a regression-based algorithm, the attributes could include the total number of variables, the specific variables to be included, and the coefficients of the variables included in the final model. A machine learning algorithm builds a model by examining many examples and attempts to find a model that minimizes "loss" or error. Model "validation" refers to comparing the model's predictions with observed data. The predictions are made on data that were excluded or "held-out" from the training process. The machine learning model's accuracy can be evaluated by comparing the model's predictions with the observed values from this "held-out" dataset. The first step of machine learning involves splitting the dataset so part of it can be used for training and the other part can be used for validation.

The dataset with obesity prevalence estimates at *time t*, obesity prevalence estimates at $time_{t-3}$, and sociodemographic data at $time_{t-3}$ for the years 2008-2016 was split twice (**Figure 5.1**). First, we split the data into two groups chronologically: (i) 2008-2013 and (ii) 2014-2016. To ensure the algorithm had enough data to "train" with, we used two-thirds of the years for model training (2008-2013) and one-third for model validation (2014-2016). Then we split neighborhoods into a training set or a validation set. Data from neighborhoods in the training set were used to build the machine learning model. Data from neighborhoods in the validation set were not used to build the machine learning model, but were used to assess the model's accuracy.

From group (i), the split consisting of 2008-2013 data, 75% of neighborhoods were randomly assigned into group (a), the training set. We used data from group (a), the training set, for the years 2008-2013 to "train" our model to predict future neighborhood-level prevalence of obesity. Once the model was "trained", we assessed the model's predictive accuracy with the "held-out" data from 2014-2016. Data from neighborhoods included in the training set, but for the years 2014-2016, were included in the model validation. Since 2008-2013 data from these neighborhoods were used to build the machine learning model, we refer to the 2014-2016 data from these neighborhoods used in the model validation process as the *in-sample* validation set. The remaining 25% of neighborhoods that were not selected into group (a), the training set, were assigned to group (b), the validation set. These neighborhoods were not used in the model training. Data from neighborhoods in the validation set for 2014-2016, referred to as the *out-of-sample validation set*, were used to validate our model and assess the accuracy and generalizability of the model predictions.

Figure 5.1. Study overview for Chapter 5



<u>Dependent variable</u>

Obesity prevalence estimates were calculated from data provided by the LA County WIC Data Mining Project (described in detail in CHAPTER 3). *Neighborhood-level childhood obesity prevalence* was defined as the percent of WIC-participating children ages 2 to 5 years who were obese during a 3-year period (e.g. 2003-2005 prevalence, 2004-2006 prevalence, etc.). A 3-year period was used to obtain stable estimates of obesity prevalence, particularly for ZIP Codes with fewer numbers of WIC children. As used in many studies, ZIP Codes were used as a proxy for neighborhoods.[111,180-183] Obesity status for children 2 to 5 years enrolled in WIC was determined using data on height, weight, and age. A child was categorized as obese if the child had a Body Mass Index (BMI) ≥95th percentile of CDC's sex-and age-specific growth reference values.[46] Neighborhood-level obesity prevalence was calculated by dividing the total number of unique children who were obese in a given ZIP Code for each 3-year period (e.g. between 2003 and 2005 for the 2005 3-year estimate) divided by the total number of unique WIC-enrolled children residing in that ZIP Code for that 3-year period. Each unique child had an obesity status for each year the child was enrolled in WIC, meaning during a 3-year period, each child could have up to 3 obesity status measures. For each 3-year period, a child had to be obese at least half of the time to be considered as obese. For example, if a child was obese in 2003, but not obese in 2004 or 2005, that child would be not counted as obese for the 2003-2005 3-year period. Due to issues of confidentiality, of the 311 ZIP Codes in LAC, only ZIP Codes with at least 30 WIC children in each 3-year period were included in the study (n= 258). *Three-year neighborhood-level childhood obesity prevalence* was calculated for 2005-2016.

Independent variables

Neighborhood-level childhood obesity prevalence from the preceding 3 years (time$_{t-3}$) was also included as an independent variable, which represents the latest obesity prevalence estimate available.

Sociodemographic data were obtained from the Census Bureau's Decennial Census and the American Community Surveys (ACS) at the level of ZIP Code Tabulation Areas (ZCTA). ZCTAs are statistical geographic representations of ZIP Codes created by aggregations of census blocks.[157,159,185] Census data are available at the ZCTA level for 2000 and 2010 onwards. For this study, we used data from the 2000 Census and 2010-2015 ACS 5-year estimates. Sociodemographic data for each ZCTA were obtained in two ways. For the years 2010-2015, the data were obtained directly from the ACS 5-year estimates for each year. We used linear interpolation to estimate neighborhood sociodemographic values of interest for 2005-2009 using data from the 2000 Census and the 2010 ACS. In this study, ZCTA-level summary data were included for the following population sociodemographic characteristics based on a review of neighborhood-level sociodemographic characteristics important for obesity: total population, population density, age, race/ethnicity, education, marital status, income, poverty status, public assistance participation, employment, place of birth, housing occupancy, age of homes, and vehicle ownership (see **Table 5.1**).[107,134-137,141,165,180,203-237] All sociodemographic variables were expressed at $time_{t-3}$, which represents the latest sociodemographic data available.

Table 5.1. Study variables for Chapter 5

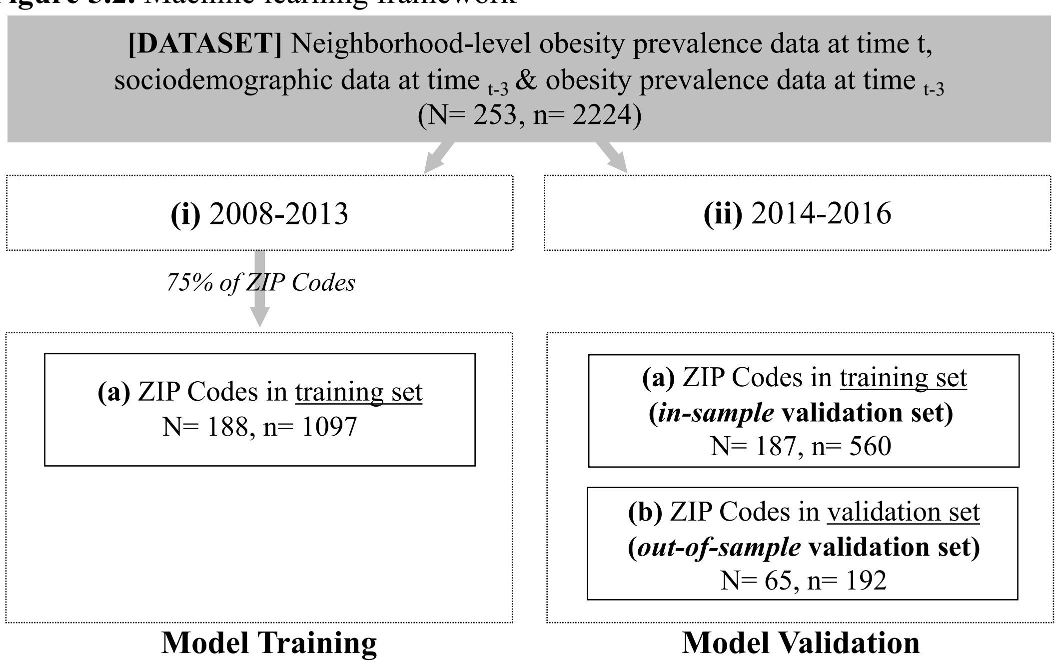
Variables	Details and categories
Total population	Total population
Population density	Persons per square mile
Age	Proportion (%) of population under 5 years Proportion (%) of dependents (under 20 and over 64) Median age
Race/Ethnicity	Proportion (%) of population: - Hispanic or Latino - Non-Hispanic Black or African American - Non white
Education	Proportion (%) of population 25 years and over - With less than a high school degree - With a Bachelor's degree or higher
Marital status	Proportion (%) of female householders, no husband present, with own children under 18 years for population 25 years and over
Income	Proportion (%) of households with income: - Less than $35,000 - $75,000 or more Median household income Per capita income
Poverty status	Proportion (%) of population with income below poverty level Proportion (%) of population with ratio of income to poverty level <2.0
Public assistance participation	Proportion (%) of households with public assistance income
Employment	Proportion (%) of civilian population 16 years and over - Not in the labor force - Unemployed - In management, professional, and related occupations
Place of birth	Proportion (%) foreign born
Housing occupancy	Proportion (%) of occupied housing units with ≤1 occupant per room Proportion (%) of vacant housing units Proportion (%) of owner-occupied housing units Average household size— Total Median value of owner-occupied housing units Proportion (%) of population living in the same house 1 year ago
Age of homes	Proportion (%) of housing units: - Built 1999 to 2000 / Built in 2005 or later[a] / Built in 2010 or later[b] - Built 1995 to 1998 / Built in 2000 to 2004[a] / Built 2000 to 2009[b] - Built 1990 to 1994 / Built 1990 to 1999[c] - Built 1980 to 1989 / Built 1980 to 1999[c] - Built 1970 to 1979 / Built 1970 to 1979[c] - Built 1960 to 1969 / Built 1960 to 1969[c] - Built 1950 to 1959 - Built 1940 to 1949 - Built 1939 or earlier
Vehicle ownership	Proportion (%) of occupied housing units with no vehicle available
Neighborhood-level obesity	Proportion (%) obese among WIC-enrolled children 2 to 5 years

[a]2010 and 2011 American Community Surveys; [b]2012, 2013, 2014, and 2015 American Community Surveys;
[c]2010, 2011, 2012, 2013, 2014, and 2015 American Community Surveys

Analysis

We linked obesity prevalence estimates at *time t* with obesity prevalence estimates from the preceding 3 years (*time* $_{t\text{-}3}$), and sociodemographic data from the preceding 3 years (*time* $_{t\text{-}3}$). The dataset was split chronologically. From the 253 unique ZIP Codes included in the dataset, 75% were randomly assigned to a training set consisting of 188 ZIP Codes and 1097 observations (in ZIP Code-years). A model to predict the prevalence of early childhood obesity was developed using data from the ZIP Codes included in the training set for the years 2008-2013. The validity of the model for predicting obesity rates was assessed by comparing the predicted obesity rates for 2014-2016 against actual rates computed using data from the in-sample validation set and the out-of-sample validation set (**Figure 5.2**).

Figure 5.2. Machine learning framework



N= number of ZIP Codes; n= number of observations (ZIP Code-years)

Model training. To train our machine learning model to predict neighborhood-level prevalence of obesity, we fit several regression models to "tune" our models. Tuning is the process of maximizing a model's predictive performance without overfitting the model. In our case, this involves the process of learning the regression coefficients of our model in order for our model to make the most accurate predictions. We used cross-validation to tune our models, which is a method for getting a reliable estimate of model performance using training data. There are several ways to cross-validate, and the most commonly used are 5- or 10-fold cross validation, with the choice depending on the size of the data available for training.

We used 5-fold cross-validation, which involved splitting the training data into 5 equal parts, or folds, to create 5 train-test splits (**Figure 5.3**). For each split, we trained our model on the 4 folds, excluding the fifth "hold-out" fold. We then evaluated how well the model performed on the "hold-out" fold. These steps were repeated 5 times, each time holding out a different fold. Finally, we computed the average performance across all 5 "hold-out" folds.

Because we had a time component to our data (yearly data), we had to ensure that there were no temporal violations when creating our splits. For example, using 2011 data to train our model, and then evaluating the model to predict 2010 obesity prevalence estimates would overestimate our model's predictive accuracy. Doing so would involve "looking ahead" into the future to predict past obesity estimates. In order to prevent violations of temporality, we performed 5-fold cross validation yearly, from 2008-2013. That is, we created 5 train-test splits for 2008, 5 train-test splits for 2009, etc. In total, we used 30 train-test splits to fit our final machine learning model, 5 train-test splits for each of the 6 years we used for training. The number of ZIP Codes and observations (in ZIP Code-years) included in the study by different aspects of the model training-testing process can be seen in **Table 5.2**.

Figure 5.3. Chronological data splits and training-validation set pairs for model training

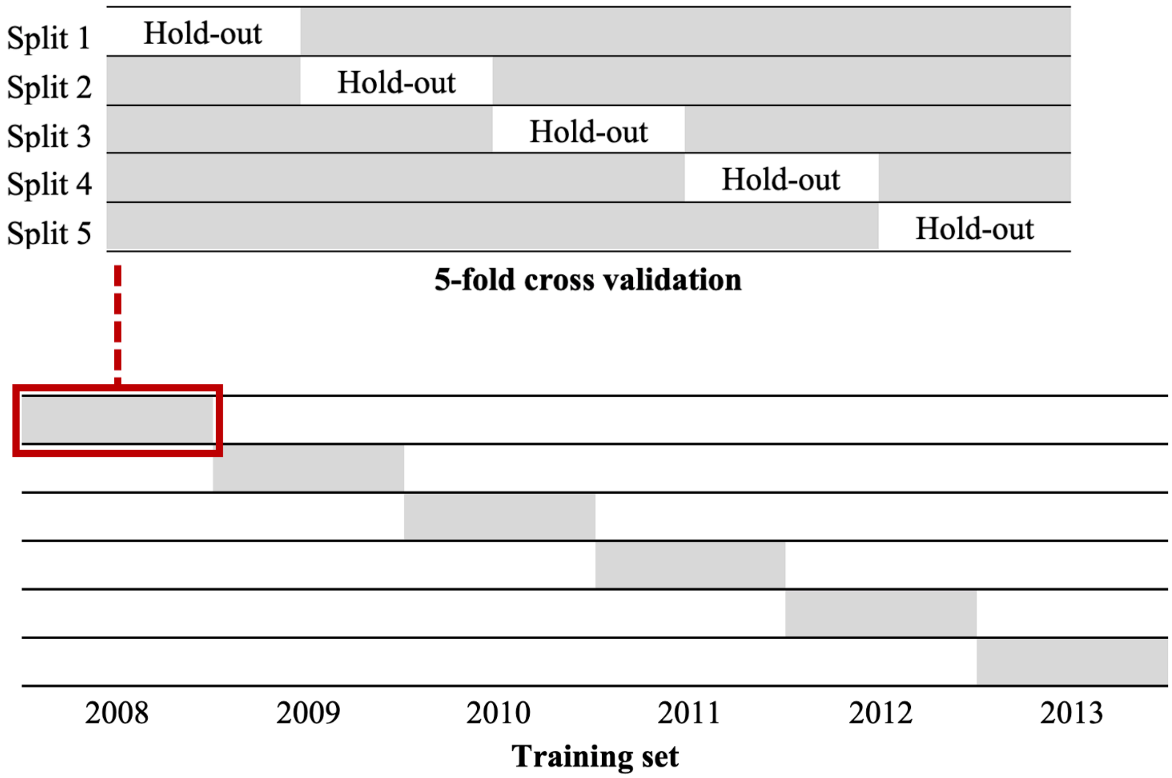


Table 5.2. Summary of ZIP Codes included by year for Chapter 5

	Overall	*Model Training*	*Model Validation*	
		Training set	In-sample validation set	Out-of-sample validation set
	N= 253, n= 2224	N= 188, n= 1097	N= 187, n= 560	N= 65, n= 192
2008	243	181		
2009	242	181		
2010	245	182		
2011	243	180		
2012	249	186		
2013	250	187		
2014	251		187	64
2015	252		187	65
2016	249		186	63

N= number of unique ZIP Codes; n= number of observations (ZIP Code-years)

To test our hypothesis that using both the latest obesity prevalence estimates and sociodemographic data available would most accurately predict future prevalence of early childhood obesity at the neighborhood level than either alone, we trained 2 separate machine learning models. To train the models, regression models in the form of (1) and (2) were fitted, respectively, using the 188 ZIP Codes and 1097 observations in the training set for 2008-2013 (see **Figure 5.2**). Model (1) included both the latest obesity prevalence estimates and sociodemographic data available, and Model (2) only included the latest sociodemographic data available. The "latest" data available were represented by 3-year time lags.

$$ObesityPrev_t = c + \beta_1(ObestityPrev_{t-3}) + \beta_i(Sociodemographic\ variables_{i_{t-3}}) \quad (1)$$

$$ObesityPrev_t = c + \beta_i\left(Sociodemographic\ variables_{i_{t-3}}\right) \quad (2)$$

Due to the large number of variables involved, the regression model was trained using least squares with LASSO (least absolute shrinkage and selection operator)[251] penalization on the regression coefficients (β), a commonly used machine learning method. LASSO has a "tuning" parameter that decides the number of variables selected in the final trained model, and this was determined by cross-validation. LASSO is a regularization method that penalizes estimators that include more covariates, especially correlated covariates, and is commonly used to reduce overfitting.[252] Regularization methods produce more parsimonious, or simple, estimators, improves generalizability, and helps to produce stable results less sensitive to small changes in estimator choices.[252]

Model validation. The trained models were validated using the chronologically separated data from (ii)(a) 2014-2016 in-sample validation set, and (ii)(b) 2014-2016 out-of-sample validation set (**Figure 5.2**).

To validate Model (1), we used information about the ZIP Codes' sociodemographic profiles and obesity prevalence estimate from the preceding 3 years, included in the 2014-2016 in-sample validation set, to predict neighborhood-level prevalence of early childhood obesity for the in-sample validation set. For example, we used the trained model to predict the prevalence of early childhood obesity for ZIP Code 90001 in 2014 using the sociodemographic profile for ZIP Code 90001 in 2011 and the obesity prevalence estimate of ZIP Code 90001 in 2011. This was repeated to predict neighborhood-level prevalence of early childhood obesity using data from the 2014-2016 out-of-sample validation set. These predicted estimates, for the in-sample and out-of-sample validation sets, were compared with the observed prevalence for each ZIP Code and year. This process was repeated to validate Model (2).

Various metrics were used to measure the degree of error between the observed and predicted results. We used the root mean square error (RMSE), the mean absolute error (MAE), and Pearson correlation (r) to measure the performance of the trained models, which are standard measures in the literature for prediction analytics.[253] Lower values of RMSE and MAE indicate better fit, as both are measures of the average magnitude of the error between predicted and observed values in the validation set. Conversely, higher values for the correlation between the predicted and observed values indicate better fit, and a more accurate model. While the results for the model validation group directly measure the model's prediction performance, it is the performance difference between the training cohort (in-sample validation set) and the validation cohort (out-of-sample validation set) that can be used to assess the degree of over-fitting in the model and its generalizability to unseen data (see **Figure 5.2**).

To examine whether the latest obesity prevalence estimate could be used as a proxy for future prevalence of early childhood obesity, we also calculated the RMSE, MAE, and Pearson correlation between the latest obesity prevalence, represented by obesity prevalence at $time_{t-3}$, with future prevalence of early childhood obesity, represented by obesity prevalence at *time t*. For example, we evaluated how well the obesity prevalence estimate in 2011 could predict, or serve as a proxy, for obesity prevalence in 2014. We compared the accuracy of this proxy with the accuracy from Model (1) and Model (2).

Statistical computing was conducted using R version 4.0.3.[190] The `caret` packaged was used for the development of the machine learning models.[254]

Results

The predictive performance of Model (1), Model (2), and (3) the proxy of the latest obesity prevalence estimates at $time_{t-3}$ predicting future obesity prevalence (at *time t*) is shown in **Table 5.3**, and as scatter plots in **Figure 5.4**. Lower values of the RMSE and the MAE indicate better fit, as both are measures of the average magnitude of the error between predicted and observed values in the validation set, whereas higher values of Pearson correlation suggest greater model accuracy.

For Model (1), the overall in-sample RMSE was 3.08 compared to 2.32 with the out-of-sample validation set. The overall MAE was slightly lower at 2.38 with the in-sample validation set compared to 2.56 with the out-of-sample validation set. The Pearson correlation between the predicted and observed values of obesity prevalence was slightly higher using the out-of-sample validation set (0.80) than the in-sample-validation set (0.79). We expect lower RMSE and MAE values resulting from the in-sample validation set compared to the out-of-sample validation set since the model had not seen any of the ZIP Codes included in the out-of-sample validation set, whereas ZIP Codes included in-sample validation set for 2008-2013 were used to train the model. The statistics from the in-sample validation reveal how well the model would hold into the future for ZIP Codes that were used to train the model. The statistics from the out-of-sample validation reveal how well the model would hold into the future generally, for ZIP codes that were not included in model training.

Model (2) underperformed compared to Model (1), as suggested by the higher RMSE and MAE values, and lower Pearson correlations. For Model (2), the overall in-sample RMSE was 3.61 compared to 3.91 with the out-of-sample validation set. The overall MAE with the in-sample validation set was 2.87, which was lower than the out-of-sample MAE of 3.22. The Pearson correlation between the predicted and observed values of obesity prevalence was 0.73 with both the in-sample and out-of-sample validation sets.

The use of the latest obesity prevalence estimates at $time_{t-3}$ as a proxy for future obesity prevalence (at *time t*) (3), outperformed Model (2), but underperformed compared to Model (1) with an RMSE of 3.53, a MAE of 2.58, and a Pearson correlation of 0.73 between observed values of obesity prevalence at $time_{t-3}$ and observed values of obesity prevalence at *time t*.

Table 5.3. Accuracy of neighborhood-level early childhood obesity prevalence estimates*

Model [1]

		Sociodemographic data *(time $_{t-3}$)* & obesity data *(time $_{t-3}$)*		
		RMSE	**MAE**	**Pearson correlation (r)**
IN-SAMPLE	**TOTAL**	**3.08**	**2.38**	**0.79**
	2014	2.71	2.09	0.81
	2015	3.29	2.49	0.76
	2016	3.22	2.55	0.81
OUT-OF-SAMPLE	**TOTAL**	**3.32**	**2.56**	**0.80**
	2014	2.94	2.32	0.84
	2015	3.27	2.54	0.86
	2016	3.71	2.84	0.72

Model [2]

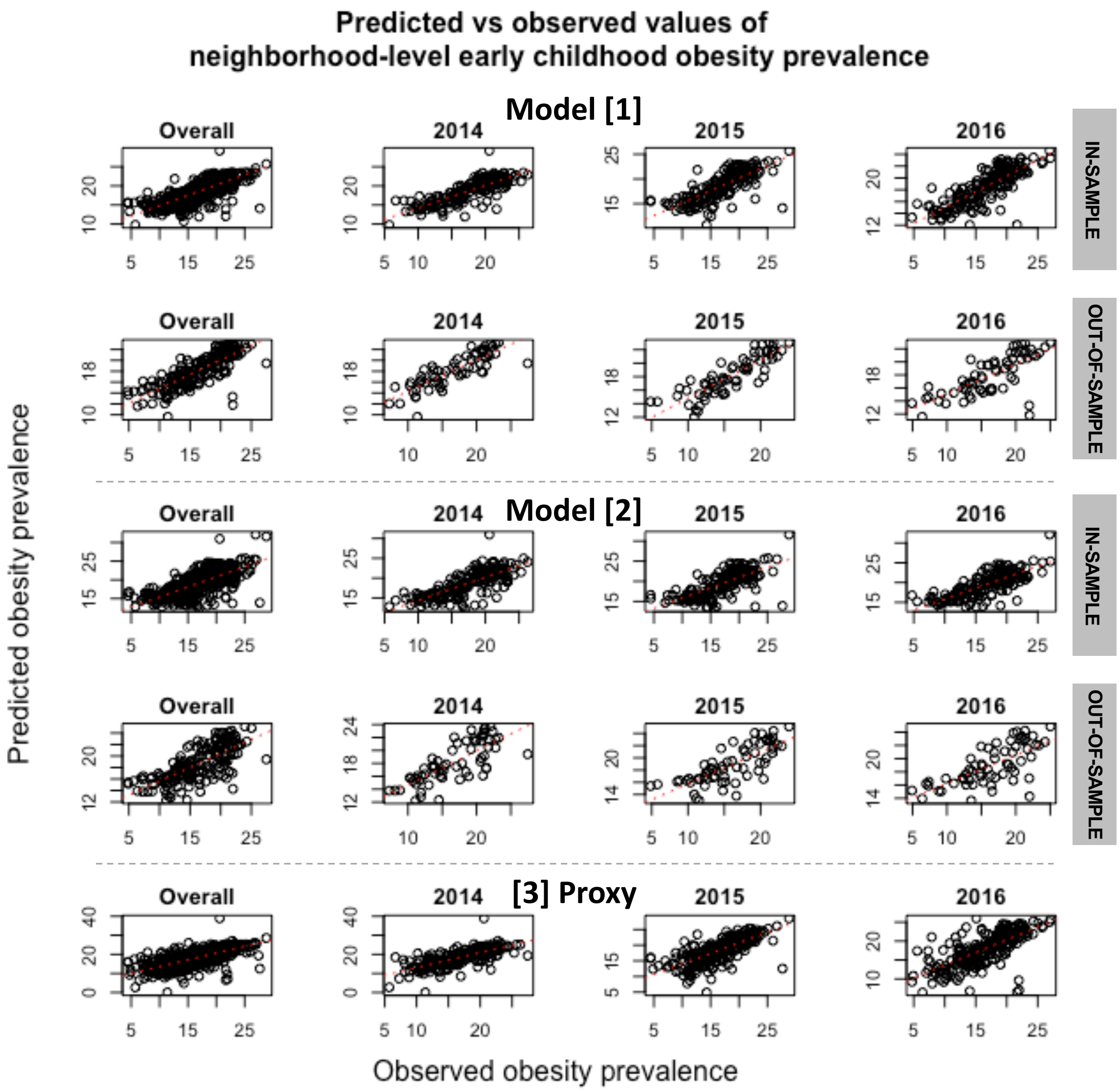
		Sociodemographic data *(time $_{t-3}$)*		
		RMSE	**MAE**	**Pearson correlation (r)**
IN-SAMPLE	**TOTAL**	**3.61**	**2.87**	**0.73**
	2014	2.97	2.30	0.76
	2015	3.86	3.04	0.71
	2016	3.93	3.27	0.78
OUT-OF-SAMPLE	**TOTAL**	**3.91**	**3.22**	**0.73**
	2014	3.30	2.83	0.76
	2015	4.08	3.35	0.76
	2016	4.29	3.47	0.70

[3] Proxy

		Obesity data *(time $_{t-3}$)*		
		RMSE	**MAE**	**Pearson correlation (r)**
OBSERVED	**TOTAL**	**3.53**	**2.58**	**0.73**
	2014	3.45	2.51	0.75
	2015	3.54	2.62	0.74
	2016	3.61	2.60	0.71

*Obesity prevalence predictions were evaluated against observed obesity prevalence from in-sample and out-of-sample validation data for Model (1) and Model (2); Observed obesity prevalence at $time_{t-3}$ was evaluated against observed obesity prevalence at time t for (3) Proxy

Figure 5.4. Scatter plots of predicted vs observed values of neighborhood-level early childhood obesity prevalence, 2014-2016



Predicted obesity prevalence estimated from Model (1) machine learning model using sociodemographic data and obesity data; (2) machine learning model using sociodemographic data; (3) obesity prevalence at $time_{t-3}$

Discussion

Estimates of both current and future childhood obesity prevalence are needed for population health planning and resource allocation. However, due to lag times between data collection and data availability, current up-to-date estimates are not available. The main objective of this study was to demonstrate the feasibility of using the most up-to-date obesity prevalence estimates and sociodemographic data available to predict future neighborhood-level prevalence of early childhood obesity using machine learning techniques.

The trained Model (1)'s predictions of obesity prevalence for 3 years of data (that were not included in the original model development), were highly correlated with the observed data, in both the ZIP Codes included in the training model (correlation 0.79) and those excluded from the development for use as a completely separated validation sample (correlation 0.80), demonstrating that the model had sufficient external validity to make good predictions, based on latest obesity prevalence estimates and sociodemographic data available alone, for ZIP Codes not included in the model development. Model (2), which included only the latest sociodemographic data available, and (3) using the latest obesity prevalence estimates at $time_{t-3}$ as a proxy for future obesity prevalence (at *time t*) provided similar estimates of future obesity prevalence, in terms of accuracy. The model-predicted and observed obesity prevalence for Model (2), as well as the observed obesity prevalence at $time_{t-3}$ and *time t* for (3), were reasonably correlated (correlation 0.73). This suggests that future obesity prevalence can be predicted just as well from the latest obesity prevalence estimate available (with a time lag) as it can with Model (2). However, Model (1) provided the most accurate estimates of future obesity prevalence, as evidenced by the higher correlation between model-predicted and obesity prevalence of obesity (0.79-0.80). As we hypothesized, the machine learning model that used both the latest obesity prevalence estimates and sociodemographic data available provided the most accurate predictions of future prevalence of early childhood obesity at the neighborhood level than either alone. Our findings suggest that the modeled estimates from machine learning models, when using both the latest obesity prevalence estimates and sociodemographic data available, may be a valid method for rapidly predicting future neighborhood-level early childhood obesity prevalence and identifying neighborhoods most burdened by early childhood obesity.

A limited number of publications have applied machine learning techniques to predict childhood obesity, though most have focused on predicting obesity at the individual level. Using existing electronic health record (EHR) data collected prior to a child's second birthday, Dugan et al[243] predicted obesity at age 2 with 85% accuracy, and Hammond et al[244] predicted obesity at age 5 with 76.1-81.7% accuracy. Using genetic profiles, Montañez et al.[245] classified individuals' susceptibility to obesity with 90.5% accuracy. Zhang et al.[246] compared logistic regression models with machine learning models to predict obesity at 3 years using data recorded before 3 years, and found that using machine learning techniques improved accuracy of prediction. Using state-level sociodemographic characteristics as predictors, Luo et al.[247] found that agreement (Pearson correlation) between the model-predicted and observed prevalence of adult obesity at the state-level was 75%. Though the application of machine learning to obesity prediction in the current literature has been applied to individual and state levels, our findings at the ZIP Code-level are similar to the published findings, with an observed Pearson correlation of 80% between model-predicted and observed prevalence of early childhood obesity.

The need to strengthen local public health data systems for surveillance at the neighborhood or small-area levels has been widely recognized, and has led to groups working towards developing small-area estimation models. Small-area estimation models generate community-specific prevalence estimates by making use of (1) the associations of obesity with individual- and community-level characteristics, (2) data from multiple years and across geographic regions, and (3) community demographic characteristics.[132] Seliske et al.[255] used hierarchical Bayesian models to estimate prevalence of obesity among adults in Census administrative units in Canada with a precision of about 83.4%. Li et al.[132] used random-effects logistic regression models to estimate prevalence of obesity in 398 communities in Massachusetts, which they considered as towns, small cities, and subdivisions of large cities, though the authors did not provide measures of validation. An advantage of machine learning techniques over more traditional methods, is it allows us to assess how accurate our model is. Machine learning techniques provide an alternative and attractive modeling method for analyzing existing data to predict future prevalence of early childhood obesity. The 80% correlation we found between model-predicted and observed prevalence of early childhood obesity is similar to the 83.4% precision found using small-area estimation modelling,[255] which requires more data and complex analytical modelling. This further highlights the utility of this sophisticated, relatively straightforward, machine learning approach to model development.

Organizations that deliver public health programs and interventions should establish need through population health assessments and surveillance, and these programs should be continuously tailored based on the needs identified. The shortage of resources for local public health programs underscores the need for up-to-date estimates of obesity prevalence across communities to support the planning, implementation, and evaluation of programs delivered to communities with the greatest needs. Estimates of both current and future childhood obesity prevalence at the neighborhood-level from machine learning models can be used to guide population health planning and resource allocation. Furthermore, analysis of situations where the measured prevalence of early childhood obesity diverge substantially from machine predictions may help to identify areas of best practice or areas with greater need for investment in action and policy to prevent and manage obesity. For example, in CHAPTER 4, we demonstrated that place-based interventions to address obesity have increasingly been implemented across LAC. These interventions may explain some of the variation in the precision of estimates.

This study had several limitations. First, the sociodemographic data obtained from the US Census Bureau are at the level of ZCTAs, while obesity prevalence estimates were calculated for ZIP Codes. A possible limitation, particularly in relation to applying the models for forward predictions of obesity prevalence, is the inability of the model to adapt to secular changes in obesity prevalence, which are likely to occur at a much faster rate than changes in the demographic profile of the regions. It is important that the results of the machine predictions are interpreted with caution, and that any extrapolation or future predictions are not stretched too far from the "training" data. It is plausible that the observed ecological associations may not be stable over time. If models trained using aggregate data estimated at one level (e.g. ZIP Codes) are to be applied to predict prevalence at a different geographic level (e.g. Service Planning Areas or counties), care must be taken to validate the approach with observed data where possible, as relationships between population demographics and NCD prevalence may differ at different geographical levels.[256] When richer data are available, more accurate prevalence estimates may be achieved.

This study had a number of strengths. First, measured heights and weights of high validity[149] were used to calculate obesity prevalence estimates. As a result, these measures provided a "gold standard" estimate of prevalence for training and testing the machine learning models. Another strength was that the demographic data used are simple, widely available, and were drawn from an entirely separate data source. Though our machine learning model that included both the latest obesity prevalence estimates and sociodemographic data available provided the best predictions, our model predicting outcomes based on sociodemographic data also provided reasonable predictions. In the absence of the availability of obesity prevalence data, the level of prediction accuracy achieved in this demonstration, could be applied to fill gaps in data collection from more traditional sources.

Conclusion

This study powerfully illustrates the strength of using existing data, particularly, the latest obesity prevalence estimates and sociodemographic data available, to predict community-level obesity burden, with the application of sophisticated modelling techniques. The research described here is a simple demonstration of the potential for machine learning techniques to contribute to the field of public health research. The technique demonstrated raises the possibility of future low-cost approaches to rapidly estimating burden of childhood obesity. In the absence of up-to-date estimates of obesity prevalence, this method provides a critical source for public health practitioners and policymakers.

CHAPTER 6: The contribution of place-based interventions to declines in neighborhood-level rates of early childhood obesity: A counterfactual approach to the evaluation of community-level interventions

Introduction

Childhood obesity is one of the most serious public health challenges of the 21st century.[4] Obesity in childhood starts very early in life obese children likelier to remain obese into adulthood, and likelier to develop non-communicable diseases like diabetes and cardiovascular diseases at a younger age.[4] In the United States (US), obesity prevalence among preschool-aged children increased dramatically over the past three decades, almost doubling between 1988-1994 and 2003-2004 from 7.2% to 13.9%.[1] While it temporarily decreased to 9.4% in 2013-2014,[1] it began to increase again to reach 13.9% in 2015-2016 (latest data available).[3] Children from low-income families are more likely to be obese than children from high income families.[49] Until recently, obesity prevalence among preschool-aged children enrolled in the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), a federal nutrition assistance program, was consistently higher than that of the overall US population of preschool-aged children. In 2004, early childhood obesity prevalence was 15.5% among WIC-enrolled children compared to 13.9% for the overall US population. In 2014, the corresponding rates for WIC-enrolled children and the overall US populations were 14.5% and 9.4% respectively suggesting a widening disparity between 2004 and 2014.[1,50] The latest available estimates suggest that this disparity may be narrowing, with obesity reaching 13.9% in 2016 (2015-2016) for both the overall US population and WIC-enrolled children of preschool age.[3,51] However, obesity rates among WIC-enrolled children may have started to rise again.

Biologically, obesity is due to energy intake that exceeds energy expenditure. Beginning in infancy, infants who are breastfed are at reduced risk of childhood obesity compared to those who are formula-fed.[257] As children transition from consuming a single food (breast milk or formula) to consuming a variety of solid foods, what and how much they eat, in addition to their levels of PA will affect their risk of obesity.[9] Taking a socioecological approach,[14] we understand that these obesity-related behaviors occur within the context of the child's home environment, which is further contextualized by the family's neighborhood environment.

Young children's diets are determined by their parents as parents determine the types of foods available at home, the portion sizes, and mealtime structures, which can either encourage or discourage children's healthy eating.[68] Parents also play the largest role in determining young children's levels of PA. For example, parents' decisions to take their child to a park to play, go for family walks, enroll their child in organized sports, or encourage their child to be active by attending or providing transportation to sporting events all influence how active young children are.[70,71] Furthermore, according to Social Cognitive Theory,[73] children's obesity-related behaviors are also shaped by what their parents do, as children learn through observation and model eating and PA behaviors of their parents.[70,71,74] However, the neighborhoods that families live in can constrain parents' choices, thereby shaping obesity-related behaviors of families and young children.

The neighborhood environment can enable healthy behaviors or hinder them, and neighborhoods indirectly affect young children's obesity risk through their parents. For example, families living in neighborhoods where large supermarkets are not easily accessible may face greater barriers to purchasing fresh fruits and vegetables.[258] In the US today, low SES families are more likely to live in such neighborhoods where not only is fresh produce not readily accessible, but unhealthy processed food sold by small corner markets is readily accessible. The neighborhood food environment largely shapes the proximity, accessibility, variety, and quality of healthy food options that parents, and therefore children, have thereby facilitating or constraining family food choices. For young children, access to safe areas to play, playground density, as well as parental perceptions of playground safety, may influence their levels of PA. For families living in poorer and more crime-ridden areas, parents may prevent their children from spending time outside, and these children may have fewer opportunities to be active. Given that young children's obesity-related behaviors are shaped by their environmental contexts, specifically, their home environment, which in turn is shaped by their neighborhood environment, approaches to health promotion have begun to shift over the past few decades from individual-based approaches that focused on individual-level behaviors to community- and population-based approaches that incorporate socioecological perspectives.[120]

In an effort to tackle obesity, communities throughout the US have engaged, to varying degrees, in creating environments that support healthy nutrition, PA, and healthy weight— the modifiable determinants of obesity.[259] Such efforts to create healthy environments to address obesity are based on both empirical evidence and social and behavioral science-based theoretical frameworks that have been used in program planning and evaluation. Place-based health promotion emphasizes population-based initiatives that are aligned with the socioecological framework[14] to address the multiple factors and levels of influence on obesity-related behaviors. Place-based multicomponent, multilevel interventions aim to influence behaviors of individuals, groups, and organizations, while also attempting to incorporate strategies to create policy and environmental changes to support individual-level behavior change.[121] Multilevel interventions address more than one level of the socioecological model[14] to target both individuals, as well as their environmental context through intervention strategies. Multicomponent interventions refer to interventions that incorporate more than one intervention strategy to achieve an improved health outcome, where an intervention strategy is defined as a plan of action that describes a method for achieving project objectives and producing defined outcomes.

Place-based interventions have shown some promise for preventing obesity, leading to modest reductions in population weight gain,[21,22,24] albeit the evidence has been mixed.[23,25] A systematic review of place-based interventions targeting children or adolescents found that six of the eight reviewed studies reported a significant improvement in at least one measure of adiposity that could be attributed to the intervention, and a meta-analysis of these six trials revealed a small reduction in Body Mass Index (BMI) z-score of 0.09 among participants in intervention communities.[21] However, these reductions in BMI z-score did not translate to reductions in the proportion of participants who were overweight or obese.[21] Another systematic review that included 51 school-based interventions found that interventions to improve weight status in preschool-aged children led to a smaller BMI increase over time among children who received the intervention relative to those that did not.[24] Conversely, a review of multilevel and multicomponent obesity-related interventions found that only three of the eight studies included in the review reported significant reductions in obesity.[23] However, four of eight studies included in the review showed significant improvement in dietary behavior and five of eight showed significant improved in levels of PA.[23] Another systematic review identifying 10 community-based interventions, only found moderate evidence to support community-based, diet–PA combined interventions that included a school component to prevent obesity, and found insufficient evidence to support community-based interventions alone.[25] Based on the published literature, evidence demonstrating the contribution of place-based interventions to changes in population-level rates of childhood obesity is very limited.[26] However, limited findings suggest that place-based interventions do lead to small reductions in population adiposity among children and adolescents, which are likely to yield important improvements in community health.[260,261]

Many organizations continue to address obesity through place-based health promotion initiatives that are multilevel, multicomponent, and implemented through multiple sectors and settings of an entire community.[21-23] Over the last decade and a half, more than $2 billion was pledged by the Centers for Disease Control and Prevention (CDC), the Robert Wood Johnson Foundation (RWJF), and The California Endowment to target and help reverse the obesity epidemic through community programs, policies, and interventions.[18-20] For example, CDC funded the Communities Putting Prevention to Work program in 2010 for 2 years in 50 communities to address obesity and tobacco use, and the Community Transformation Grants (CTGs) in 2011, provided $103 million in CTGs to 61 state and local government agencies, tribes and territories, and non-profit organizations in 36 states to implement community-level interventions.[262] In 2007, RWJF pledged $500 million to fund efforts to help reverse the childhood obesity epidemic, which included Healthy Kids, Healthy Communities, a national program to implement healthy eating and active living initiatives focusing on children who are at greatest risk for obesity.[139] In 2010, The California Endowment pledged $1 billion to build healthy communities to improve the health and well-being of its residents.[18]

Considerable investments have been also been made to promote healthy communities for families in Los Angeles County (LAC).[151,175,176] LAC is most populous county in the US with an ethnically diverse population of over 10 million residents, and a high percentage of children who are low-income.[52] In LAC, over half of all infants and children under 5 years are enrolled in WIC.[147] In LAC, large-scaled local initiatives to address obesity began in about 2005 when several government, private, and health systems organizations began implementing place-based obesity prevention efforts.[151,175-177] In CHAPTER 4, we showed that obesity neighborhood-level prevalence of obesity among WIC-enrolled children 2 to 5 years peaked at 20.38% in 2010, then decreased to 17.79% in 2015. Communities across the US, including communities in LAC, have implemented programs and policies designed to reduce childhood obesity, but the characteristics of these programs and policies that may have played a role in the stabilization of obesity rates is unclear.[263]

Place-based interventions, especially interventions that are multilevel and multicomponent, are difficult to evaluate. The "gold standard" for evaluating community-level interventions is the cluster randomized trial (CRT), where communities, rather than individuals, are randomized into an intervention or a comparison group. However, CRTs are often expensive, challenging to implement, and it is unlikely for communities to agree to be randomly assigned to a comparison group. It is also impractical and often inappropriate to implement a "one size fits all" intervention for all communities given that communities have different needs. Due to these challenges, many place-based interventions are not CRTs. Another obstacle to evaluating place-based interventions is how best to correctly measure or quantify the intervention or interventions that a given community received. With the influx of funding to address obesity at the community level, many communities now receive a variety of interventions simultaneously, as various organizations implement interventions without coordination. Without identifying all of the interventions that a community received at a given time, we are unable to fully understand how the various place-based interventions contributed to any changes in obesity prevalence seen at the community-level.

To address these challenges, we need to identify a good comparison group of communities that did not receive obesity-related interventions. To understand and quantify the effectiveness of obesity-related interventions, we would like to quantify the difference between the obesity prevalence of a neighborhood that received obesity-related interventions in a given year and the obesity prevalence of that same neighborhood if that neighborhood had not received obesity-related interventions in that same year. To illustrate, in 2005, Neighborhood Z received (or was "exposed" to) obesity-related interventions and had an obesity prevalence of 20%. We would like to know what the obesity prevalence would have been in Neighborhood Z in 2005 if that neighborhood had not received (or was "unexposed" to) obesity-related interventions— this is referred to as the "counterfactual" outcome. If Neighborhood Z had not received any obesity-related interventions in 2005, which is counter to the fact, we would hypothesize that the counterfactual obesity prevalence of Neighborhood Z would be higher than 20%. We may be able to leverage on the predictive capabilities of machine learning techniques to create a hypothetical comparison group of neighborhoods that did not receive obesity-related interventions. Doing so has the potential to push evaluation research forward.

Place-based interventions have been found to lead to small reductions in population adiposity,[21,22,24] but do place-based interventions contribute to declines in population-level rates of childhood obesity? If place-based interventions do reduce population-level rates of childhood obesity, which types of interventions produce the greatest reductions in obesity? We hypothesize that neighborhoods that received more place-based interventions would see greater declines in neighborhood-level prevalence of childhood obesity. We also hypothesize that neighborhoods that received (i) multicomponent interventions, (ii) multilevel interventions, and (iii) interventions addressing more obesity-related behaviors— that is interventions addressing PA, diet, and breastfeeding (BF), would see greater declines in neighborhood-level prevalence. Our primary objectives are to: (1) determine whether place-based interventions have contributed to declines in early childhood obesity prevalence among 2- to 5-year-old WIC-enrolled children residing in LAC, and (2) identify which types of interventions produced the greatest reductions in neighborhood-level prevalence of early childhood obesity. To achieve these objectives, we will apply emerging machine learning methods.

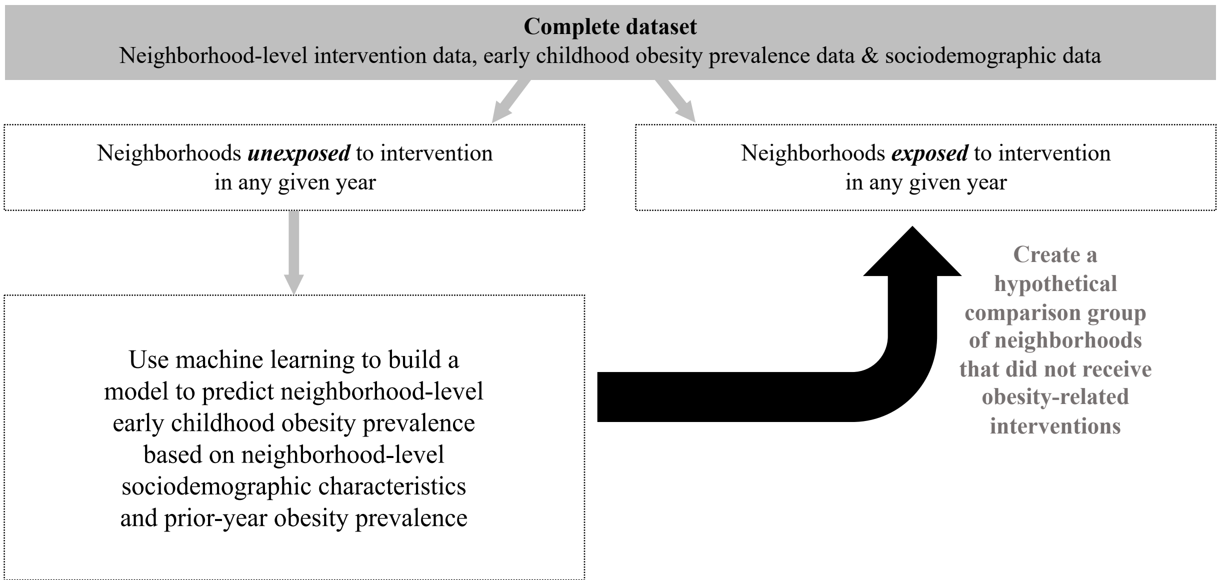
Methods

Overview

We applied machine learning methods to create a hypothetical comparison group of neighborhoods that did not receive obesity-related interventions. To do this, we linked neighborhood-level intervention data gathered by ECOSyS with neighborhood-level early childhood obesity prevalence data from the WIC Data Mining Project, and neighborhood-level sociodemographic data from the Decennial Census and ACS (referred to as the complete dataset). We then identified neighborhoods that did not receive any obesity-related intervention in a given year ("unexposed" to intervention). Machine learning techniques[264] were applied to these neighborhoods to develop a prediction model. This prediction model was subsequently applied to those neighborhoods that received interventions in a given year to estimate what obesity prevalence would have been in these neighborhoods had they not received interventions. This approach allowed us to evaluate the effect of place-based interventions without actual comparison neighborhoods by estimating the change in obesity prevalence (with versus without interventions).

To prepare our data for machine learning, we split the dataset into a model training or a model validation group. First, we split the complete dataset into two groups: 1) neighborhoods unexposed to obesity-related interventions in a given year, and 2) neighborhoods exposed to obesity-related interventions in a given year. Using data from the first group (neighborhoods unexposed to intervention), we used machine learning to build a model to predict neighborhood-level prevalence of early childhood obesity based on neighborhood-level sociodemographic characteristics and prior-year obesity prevalence. To create a hypothetical comparison group of neighborhoods that did not receive obesity-related interventions, we applied the machine learning model to data from the second group (neighborhoods exposed to intervention). This allowed us to estimate the counterfactual early childhood obesity prevalence for neighborhoods exposed to intervention. That is, for neighborhoods exposed to intervention, we were able to estimate what their obesity prevalence would have been under no intervention (**Figure 6.1**).

Figure 6.1. Study overview for Chapter 6



Study variables

Intervention data were obtained from the ECOSyS Intervention Database (described in detail in CHAPTER 3). All obesity-related interventions that neighborhoods received from major obesity-related initiatives in LAC in a given year were quantified by ZIP code. We used ZIP Codes as a proxy for neighborhoods, as it was the smallest geographic unit for which we were able to extract intervention data.

In the ECOSyS Intervention Database, each obesity-related program implemented in LAC was described by one or more projects and the intervention strategies used by each project. These intervention strategies were categorized using a modified typology that identified 10 strategies, expanding on the 9 strategies that were originally identified by ECOSyS.[150,154] Each of the 10 intervention strategies was broadly categorized as macro-level or micro-level (**Table 6.1**).

A macro-level strategy is defined as one that does not directly target individuals but may affect the larger community. The 4 macro-level intervention strategies identified were: (i) government policies, (ii) public institutional policies, (iii) infrastructure investments, and (iv) business practices. A micro-level strategy is defined as one that directly targets individuals to modify obesity-related behaviors. The 6 micro-level intervention strategies identified were: (i) group education, (ii) counseling, (iii) health communication & social marketing, (iv) home visitation, (v) screening & referral, and (vi) staff training. In the original typology of childhood obesity intervention strategies[150,154] staff training was grouped together with group education. For each project, we also identified the obesity-related behaviors addressed by the intervention, specifically, PA, diet, and/or breastfeeding (BF).

Table 6.1. Typology of intervention strategies

Type	Intervention strategies
Macro-level strategies[1]	Government policies
	Public institutional policies
	Infrastructure investments
	Business practices
Micro-level strategies[2]	Group education
	Counseling
	Health communication & social marketing
	Home visitation
	Screening & referral
	Staff training*

Typology modified from Wang et al. 2018[150]
[1]Strategies that indirectly affect the larger community and obesity-related behaviors and practices
[2]Strategies that directly target individuals to modify obesity-related behaviors
*Originally grouped together with the counseling[150]

We used ***intervention strategy count*** to quantify the 'amount' of intervention that a neighborhood (ZIP Code) received, which was operationalized as the sum total of intervention strategies implemented in each ZIP Code in a given year. ***Total types of intervention strategies*** a neighborhood received was operationalized as the sum total of the different types of intervention strategies implemented in each ZIP Code in a given year. Since there are 10 total intervention strategies (**Table 6.1**), the total types of intervention strategies a ZIP Code could receive in a given year ranged from 1 to 10.

We also identified the level of the socioecologic model[14] that interventions addressed, and determined whether or not neighborhoods received any macro-level intervention or any micro-level intervention. A neighborhood that received any of the 4 macro-level intervention strategies in a given year was considered to have received any ***macro-level intervention***. Similarly, a neighborhood that received any of the 6 micro-level intervention strategies in a given year was considered to have received any ***micro-level intervention***. ***Total types of macro-level intervention strategies*** and ***total types of micro-level intervention strategies*** a neighborhood received were operationalized as the sum total of the different types of macro-level and micro-level intervention strategies, respectively, implemented in each ZIP Code in a given year. Since there are 4 macro-level intervention strategies and 6 micro-level intervention strategies, the total types of macro-level and micro-level intervention strategies a ZIP Code could receive in a given year ranged from 0 to 4 and 0 to 6, respectively.

The obesity-related behaviors addressed by interventions were ***PA***, ***diet***, and ***BF***. For example, a ZIP Code that received any intervention addressing PA in a given year would be categorized as having received a PA intervention (quantified as a binary variable).

Using data from the LA County WIC Data Mining Project (described in detail in CHAPTER 3), ***neighborhood-level childhood obesity prevalence*** was defined as the percent of WIC-participating children ages 2 to 5 years who were obese during a 3-year period (e.g. 2003-2005 prevalence, 2004-2006 prevalence, etc.). A 3-year period was used to obtain stable estimates of obesity prevalence, particularly for ZIP Codes with fewer WIC children. Obesity status for children 2 to 5 years enrolled in WIC was determined using data on height, weight, and age. A child was categorized as obese if the child had a body mass index (BMI) ≥95th percentile of CDC's sex- and age-specific growth reference values.[46] Neighborhood-level obesity prevalence was calculated by dividing the total number of unique children who were obese in a given ZIP Code for each 3-year period (e.g. between 2003 and 2005 for the 2005 3-year estimate) divided by the total number of unique WIC-enrolled children residing in that ZIP Code for that 3-year period. Each unique child had an obesity status for each year the child was enrolled in WIC, meaning during a 3-year period, each child could have up to 3 obesity status measures. For each 3-year period, a child had to be obese at least half of the time to be considered as obese. For example, if a child was obese in 2003, but not obese in 2004 or 2005, that child would be not be counted as obese for the 2003-2005 3-year period. Due to issues of confidentiality, of the 311 ZIP Codes in LAC, only ZIP Codes with at least 30 WIC children in each 3-year period were included in the study (n= 258).

To build our machine learning model, we used neighborhood-level sociodemographic data and prior-year neighborhood-level prevalence of early childhood obesity to predict neighborhood-level prevalence of early childhood obesity. Census data are available at the ZCTA level for 2000 and 2010 onwards. For this study, we used the 2000 Census, and the ACS

5-year estimates for 2010-2015. Since data are not available for 2005-2009, we used linear interpolation to estimate neighborhood sociodemographic values of interest for 2005-2009 using data from the 2000 Census and the 2010 ACS. In CHAPTER 5, we found that neighborhood-level sociodemographic characteristics could reasonably predict neighborhood-level early childhood obesity prevalence using a dataset that included sociodemographic characteristics known to be associated with neighborhood-level obesity prevalence.[107,134-137,141,165,180,203-237] We used those same sociodemographic variables for this study, which included the following ZIP Code-level summary data: total population, population density, age, race/ethnicity, education, marital status, income, poverty status, public assistance participation, employment, place of birth, housing occupancy, age of homes, and vehicle ownership (**Table 6.2**).

Table 6.2. Study variables for Chapter 6

Variables	Details and categories
Total population	Total population
Population density	Persons per square mile
Age	Proportion (%) of population under 5 years Proportion (%) of dependents (under 20 and over 64) Median age
Race/Ethnicity	Proportion (%) of population: - Hispanic or Latino - Non-Hispanic Black or African American - Non white
Education	Proportion (%) of population 25 years and over - With less than a high school degree - With a Bachelor's degree or higher
Marital status	Proportion (%) of female householders, no husband present, with own children under 18 years for population 25 years and over
Income	Proportion (%) of households with income: - Less than $35,000 - $75,000 or more Median household income Per capita income
Poverty status	Proportion (%) of population with income below poverty level Proportion (%) of population with ratio of income to poverty level <2.0
Public assistance participation	Proportion (%) of households with public assistance income
Employment	Proportion (%) of civilian population 16 years and over - Not in the labor force - Unemployed - In management, professional, and related occupations
Place of birth	Proportion (%) foreign born
Housing occupancy	Proportion (%) of occupied housing units with ≤1 occupant per room Proportion (%) of vacant housing units Proportion (%) of owner-occupied housing units Average household size— Total Median value of owner-occupied housing units Proportion (%) of population living in the same house 1 year ago
Age of homes	Proportion (%) of housing units: - Built 1999 to 2000 / Built in 2005 or later[a] / Built in 2010 or later[b] - Built 1995 to 1998 / Built in 2000 to 2004[a] / Built 2000 to 2009[b] - Built 1990 to 1994 / Built 1990 to 1999[c] - Built 1980 to 1989 / Built 1980 to 1999[c] - Built 1970 to 1979 / Built 1970 to 1979[c] - Built 1960 to 1969 / Built 1960 to 1969[c] - Built 1950 to 1959 - Built 1940 to 1949 - Built 1939 or earlier
Vehicle ownership	Proportion (%) of occupied housing units with no vehicle available
Neighborhood-level obesity	Proportion (%) obese among WIC-enrolled children 2 to 5 years

[a]2010 and 2011 American Community Surveys; [b]2012, 2013, 2014, and 2015 American Community Surveys;
[c]2010, 2011, 2012, 2013, 2014, and 2015 American Community Surveys

We used the trained machine learning model to predict ***expected obesity prevalence without intervention*** for neighborhoods exposed to intervention in a given year, yearly. ***Observed obesity prevalence with intervention*** was operationalized as *neighborhood-level childhood obesity prevalence* for neighborhoods exposed to intervention, which we described above. To calculate our outcome of interest, ***change in obesity prevalence***, we subtracted *expected obesity prevalence without intervention* from *observed obesity prevalence with intervention* (**Eq 1**). Negative values of *change in obesity prevalence* suggest that the interventions had a beneficial effect on obesity prevalence.

Eq 1.

$$\begin{aligned} &Change\ in\ obesity\ prevalence \\ &\qquad = Observed\ obesity\ prevalence\ with\ intervention \\ &\qquad - Expected\ obesity\ prevalence\ without\ intervention \end{aligned}$$

Statistical modeling and analyses

Data on community-level interventions addressing obesity implemented in LAC were linked with neighborhood-level sociodemographic data, and neighborhood-level prevalence of early childhood obesity by ZIP Code and year for 2006-2015 (referred to as the complete dataset).

To use machine learning to create a hypothetical comparison group of neighborhoods that did not receive obesity-related interventions, we determined the intervention status of each ZIP Code for each year between 2006 and 2015, and assigned the ZIP Code to either the unexposed subset or exposed subset for the relevant year (**Figure 6.2**). For example, if ZIP Code 90001 was unexposed to intervention in 2006 and 2007, but was exposed to intervention in 2008 onwards, ZIP Code 90001 would be included in the unexposed subset for the years 2006 and 2007, and the exposed subset for the years 2008-2015. To ensure a large enough sample size to build a machine learning model, we identified and removed interventions that targeted all neighborhoods in LAC, assuming that any county-wide intervention would have the same effect on each neighborhood it reached or targeted.

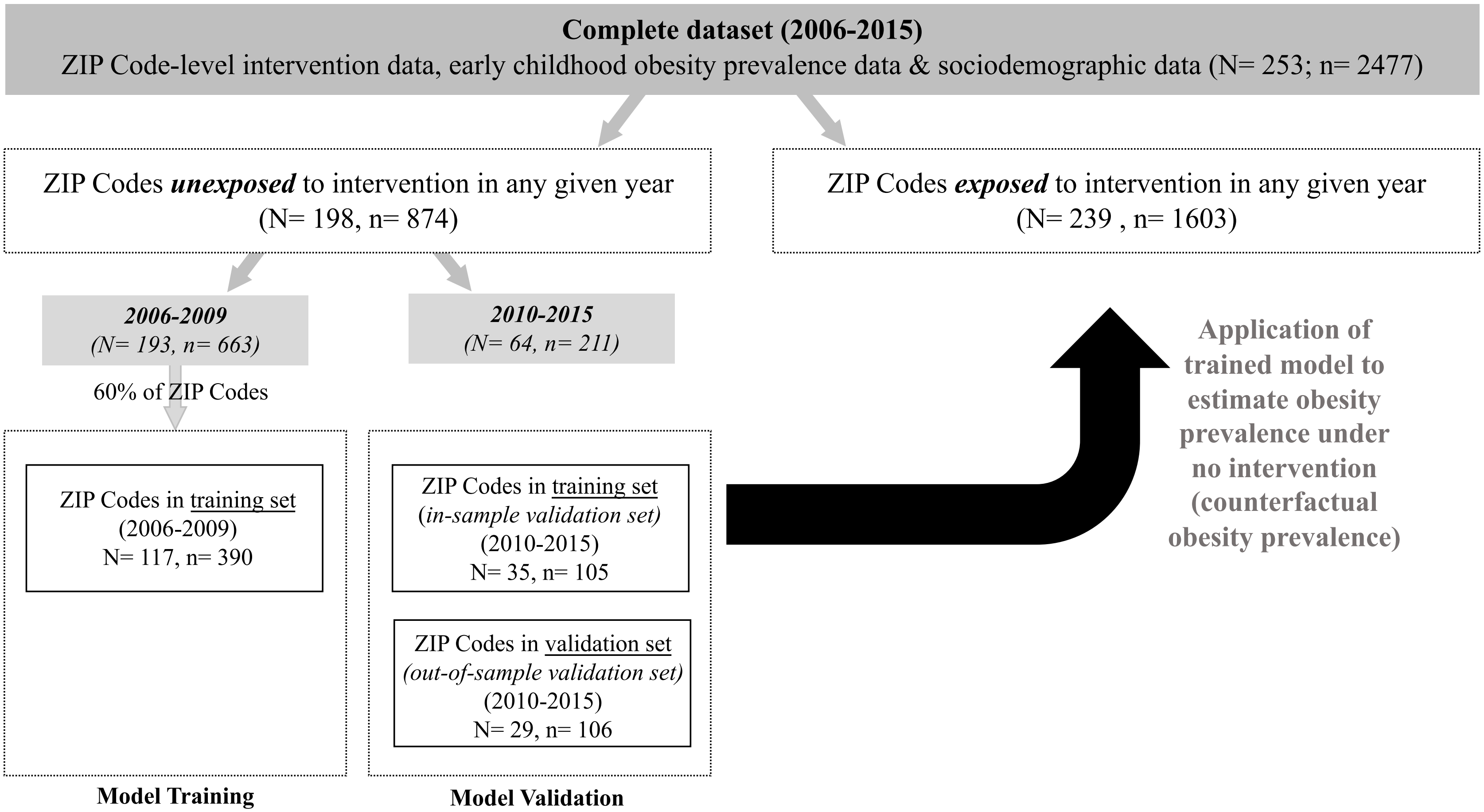
To train and validate our model, we split the data from the unexposed subset chronologically into 2 groups: (i) 2006-2009 and (ii) 2010-2015. In CHAPTER 4, we showed that the number of place-based interventions implemented in LAC began to increase around 2010. Using the unexposed subset of the complete dataset with 193 unique ZIP Codes for 2006-2009, 60% of the ZIP Codes (N= 117) were randomly assigned to the training set. Overall, 117 unique ZIP Codes were randomized into this training set, with the total number of ZIP Codes unexposed to intervention varying by year (**Table 6.3**). This training set for 2006-2009 consisting of 117 ZIP Codes and 390 observations (in ZIP Code-years) from 2006-2009 was used to train the machine learning model to predict the neighborhood-level prevalence of early childhood obesity among WIC-enrolled children aged 2 to 5 years by fitting regression models in the form of **Eq 2** (see **Figure 6.2**).

Eq 2.

$$ObesityPrev_t = c + \beta_i(Sociodemographic\ variables_{i_t}) + \beta(ObesityPrev_{t-1})$$

We trained the model using elastic net,[265] which is a commonly used regularization method used in machine learning. Regularization involves penalizing estimators that include more covariates or predictors, especially correlated covariates. Regularization methods, such as elastic net, produce more parsimonious (or simple) models, improves generalizability, and helps produce stable results that are less sensitive to small changes in covariate choices.[252] Because we had a time component to our data (yearly data), to prevent violations of temporality, we performed 5-fold cross validation yearly, from 2006-2009. That is, we created 5 train-test splits for 2006, 5 train-test splits for 2007, etc. In total, we used 20 train-test splits to fit our final machine learning model, 5 train-test splits for each of the 4 years we used for training.

Figure 6.2. Framework for analysis



N= number of ZIP Codes; n= number of observations (ZIP Code-years)

The chronologically separated data for 2010-2015 were used to test the accuracy of the model's predictions and therefore not used in the model development. The data from 2010-2015 was split into two groups: (a) the in-sample training set consisting of the ZIP Codes used for model training, and (b) the out-of-sample validation set consisting of ZIP Codes that were not used for model training. The total number of ZIP Codes used in this machine learning process can be seen in **Table 6.3**.

ZIP Codes in the training set for the years 2006-2009 were used for model training, whereas the same ZIP Codes in the training set for the years 2010-2015 were not involved in the model training process. To evaluate the model's performance, we calculated the difference between the obesity prevalence values predicted by our model and the actual observed obesity prevalence values in the in-sample training set and the out-of-sample validation set, respectively.

Table 6.3. Summary of ZIP Codes ***unexposed*** to intervention by year

	Overall	***Model Training***		
		Training set		
	N= 193, n= 874	N= 117, n= 409		
2006	186	113		
2007	185	114		
2008	181	112		
2009	111	70		
			Model Validation	
			In-sample training set	Out-of-sample validation set
			N= 35, n= 105	N= 29, n= 106
2010	33		16	17
2011	31		15	16
2012	28		13	15
2013	51		28	23
2014	36		18	18
2015	32		15	17

N= number of unique ZIP Codes; n= number of observations (ZIP Code-years)

After we evaluated the model's performance, we applied the trained model to the ZIP Codes that were exposed to obesity-related interventions between 2006 and 2015. This allowed us to estimate the *expected obesity prevalence without intervention*, that is, what each neighborhood's obesity prevalence would have been if it had not received any interventions (counterfactual obesity prevalence). This process allowed us to estimate the *change in obesity prevalence* between neighborhoods exposed to interventions (observed) and neighborhoods unexposed to interventions (predicted from machine learning model).

To evaluate whether place-based interventions contribute to declines in population-level rates of early childhood obesity among WIC-enrolled children 2 to 5 years in LAC, we tested the hypothesis that neighborhoods that received more place-based interventions would see greater declines in neighborhood-level prevalence of childhood obesity by running a fixed-effects linear model of *change in obesity prevalence* on prior-year *intervention strategy count* (**Eq 3**). We assumed a 1-year time lag of intervention effects.

Eq 3.

$$Change\ in\ ObesityPrev_{it} = \alpha_i + \beta_1 IntStrategyCount_{i(t-1)} + \mu_{it}$$

To identify which types of interventions produced the greatest reductions in neighborhood-level early childhood obesity, we ran several fixed-effects linear models. To test the hypothesis that neighborhoods that received multicomponent interventions would see greater declines in neighborhood-level prevalence of obesity, we ran a fixed-effects linear model of *change in obesity prevalence* on *total types of intervention strategies* neighborhoods received in the previous year, adjusting for prior-year intervention strategy count (**Eq 4**).

Eq 4.

$$Change\ in\ ObesityPrev_{it} = \alpha_i + \beta_1 TotalTypesIntStrategies_{i(t-1)} + \beta_2 IntStrategyCount_{i(t-1)} + \mu_{it}$$

To test the hypothesis that neighborhoods that received multilevel interventions would see greater declines in neighborhood-level obesity prevalence, we ran a fixed-effects linear model of *change in obesity prevalence* on receipt of any *macro-level intervention* in the previous year, receipt of any *micro-level intervention* in the previous year, and receipt of both *macro-* and *micro-level interventions* in the previous year, and adjusted for prior-year intervention strategy count (**Eq 5**).

Eq 5.

$$\begin{aligned} Change\ in\ ObesityPrev_{it} &= \alpha_i + \beta_1 MacroInt_{i(t-1)} + \beta_2 MicroInt_{i(t-1)} + \beta_3 MacroMicroInt_{i(t-1)} \\ &+ \beta_4 IntStrategyCount_{i(t-1)} + \mu_{it} \end{aligned}$$

To examine whether macro-level or micro-level intervention strategies were more effective in reducing neighborhood-level obesity prevalence, we ran a fixed-effects linear model of *change in obesity prevalence* on *total types of macro-level intervention strategies* and *total types of micro-level intervention strategies* that neighborhoods received in the previous year, adjusting for prior-year intervention strategy count (**Eq 6**).

Eq 6.

$$\begin{aligned} Change\ in\ ObesityPrev_{it} &= \alpha_i + \beta_1 TotTypesMacroIntStrategies_{i(t-1)} \\ &+ \beta_2 TotTypesMicroIntStrategies_{i(t-1)} + \beta_3 IntStrategyCount_{i(t-1)} + \mu_{it} \end{aligned}$$

To test the hypothesis that neighborhoods that received interventions addressing more obesity-related behaviors would see greater declines in neighborhood-level prevalence, we ran a fixed-effects model of *change in obesity prevalence* on receipt of interventions addressing *PA* in the previous year, *diet* in the previous year, and *BF* in the prior 3 years, both *PA* and *diet*, both *PA* and *BF*, both *diet* and *BF*, and all 3 behaviors (*PA*, *diet*, and *BF*), adjusting for prior-year intervention strategy count (**Eq 7**). Since we are estimating change in obesity prevalence for children 2 to 5 years, breastfeeding interventions had to have occurred around the time when children would have been breastfed. We used a 3-year lag in BF interventions to account for this.

Eq 7.

$$\begin{aligned}
&Change\ in\ ObesityPrev_t \\
&\qquad = \alpha_i + \beta_1 PA\ Int_{i(t-1)} + \beta_2 Diet\ Int_{i(t-1)} + \beta_3 BF\ Int_{i(t-3)} \\
&\qquad + \beta_4 (PA\ Int_{t-1} \cdot Diet\ Int_{t-1})_i + \beta_5 (PA\ Int_{t-1} \cdot BF\ Int_{t-3})_i \\
&\qquad + \beta_6 (Diet\ Int_{t-1} \cdot BF\ Int_{t-3})_i + \beta_7 (PA\ Int_{t-1} \cdot Diet\ Int_{t-1} \cdot BF\ Int_{t-3})_i \\
&\qquad + \beta_8 IntStrategyCount_{i(t-1)}
\end{aligned}$$

Statistical computing was conducted using R version 4.0.3.[190] The `caret` packaged was used for the development of the machine learning models.[254]

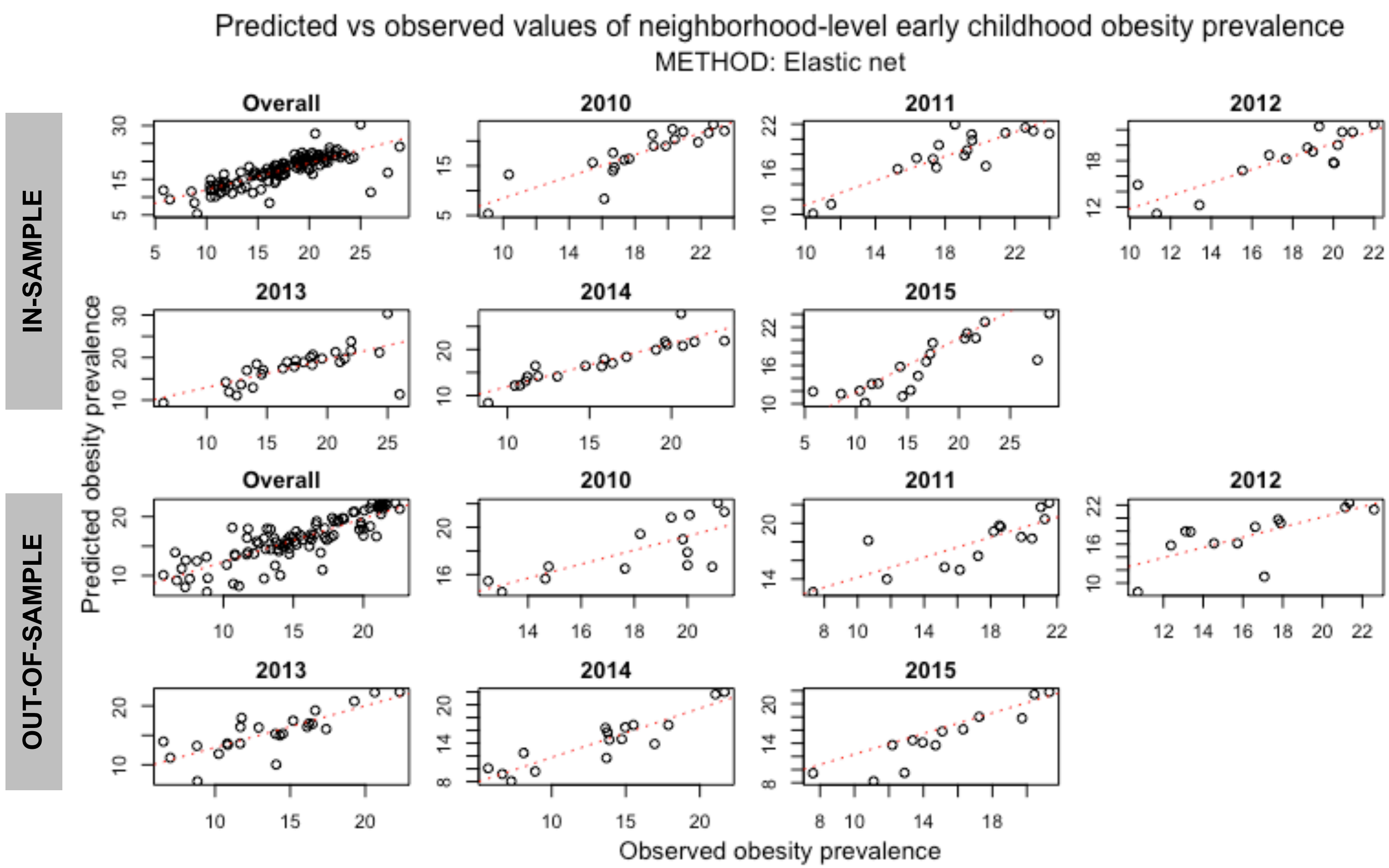
Results

Our first objective was to use machine learning to create a hypothetical comparison group of neighborhoods that did not receive obesity-related interventions. The performance of the trained prediction model on both the in-sample training set and the out-of-sample validation set is shown in **Table 6.4**, and as scatter plots in **Figure 6.3**. Lower values of the root mean square error (RMSE) and the mean absolute error (MAE) indicate better fit, as both are measures of the average magnitude of the error between predicted and observed values in the validation set. The RMSE was 2.77 with the in-sample training set compared to 2.54 with the out-of-sample validation set, suggesting that our model predicted obesity prevalence better for ZIP Codes that were not included in training the model. The MAE was slightly lower with the in-sample validation set at 1.87 compared to 1.96 with the out-of-sample validation set. The trained model's predictions of obesity prevalence for 6 years of data (that were not included in the original model development), were highly correlated with the observed data, in both the ZIP Codes included in the training model (correlation 0.80) and those excluded from the development for use as a completely separated validation sample (correlation 0.84), demonstrating that the model had sufficient external validity to make good predictions, based on sociodemographic data and prior-year obesity prevalence, for ZIP Codes not included in the model development.

Table 6.4. Accuracy of early childhood obesity estimates evaluated using in-sample training set data from 2010-2015 and out-of-sample validation set data from 2010- 2015

		RMSE	MAE	R
IN-SAMPLE	**TOTAL**	**2.77**	**1.87**	**0.80**
	2010	2.48	1.79	0.87
	2011	1.75	1.34	0.88
	2012	1.84	1.42	0.87
	2013	3.55	2.26	0.69
	2014	2.40	1.78	0.91
	2015	3.45	2.34	0.83
OUT-OF-SAMPLE	**TOTAL**	**2.54**	**1.96**	**0.84**
	2010	2.00	1.68	0.74
	2011	2.71	1.84	0.81
	2012	2.94	2.39	0.72
	2013	3.06	2.43	0.80
	2014	2.19	1.75	0.91
	2015	1.62	1.31	0.92

Figure 6.3. Scatter plots of predicted vs observed values of neighborhood-level early childhood obesity prevalence, 2010-2015



A summary of the mean *change in obesity prevalence*, that is, the difference between observed obesity prevalence with intervention and model-predicted obesity prevalence without intervention, by year can be seen in **Table 6.5**. Negative values indicate that neighborhoods that received obesity-related interventions had a lower mean prevalence of early childhood obesity than would be expected if those same neighborhoods had not received any intervention. In 2006 and 2007, obesity prevalence in neighborhoods that received interventions was, on average, 0.27% points (95% confidence interval (CI): 0.05, 0.48) and 0.33% points (95% CI: 0.01, 0.65) higher than would be expected if those neighborhoods had not received any intervention. Beginning in 2008, this trend reversed. Neighborhoods that received interventions had a mean obesity prevalence that was -0.17% points (95% CI: -0.42, 0.08) than expected if those neighborhoods had not received any intervention. Obesity prevalence in neighborhoods that received interventions continued to steadily decrease over time compared to what would have been expected if those neighborhoods had not received any intervention. In 2015, obesity prevalence in neighborhoods that received intervention was 1.20% points (95% CI: 0.98, 1.42) lower than would be expected if those neighborhoods had not received any intervention.

Table 6.5. Summary of mean change in obesity prevalence by year, 2006-2015

Year	N	Change in obesity prevalence*
		Mean (95% CI)
2006	54	0.27 (0.05, 0.48)
2007	58	0.33 (0.01, 0.65)
2008	62	-0.17 (-0.42, 0.08)
2009	136	-0.34 (-0.60, -0.08)
2010	215	-0.33 (-0.55, -0.11)
2011	219	-0.48 (-0.70, -0.25)
2012	223	-0.42 (-0.66, -0.18)
2013	201	-0.92 (-1.12, -0.73)
2014	215	-1.05 (-1.26, -0.83)
2015	220	-1.20 (-1.42, -0.98)

N= Number of ZIP Codes that received interventions
*Refers to the difference between observed obesity prevalence with intervention and model-predicted obesity prevalence without intervention

<u>Do place-based interventions contribute to declines in population-level rates of early childhood obesity?</u>

We tested the hypothesis that neighborhoods that received more place-based interventions would see greater declines in neighborhood-level prevalence of childhood obesity. The results from the fixed-effects model of *change in obesity prevalence* on prior-year *intervention strategy count* can be seen in **Table 6.6**. For each additional intervention strategy neighborhoods received in the previous year, obesity prevalence would be expected to be 0.04% points (95% CI: 0.03, 0.06) lower compared to if those neighborhoods had not received any intervention. A neighborhood that received 30 intervention strategies in the previous year would, for example, be expected to have a prevalence of early childhood obesity that was 1.20% points lower compared to if that neighborhood had not received any obesity-related interventions.

Table 6.6. Fixed-effects model of the change in obesity prevalence on prior-year intervention strategy count

	Fixed-effects model, Change in obesity prevalence* β (95% CI)
Intervention strategy count$_{t-1}$	**-0.04 (-0.06, -0.03)**

*Refers to the difference between observed obesity prevalence with intervention and expected obesity prevalence without intervention

<u>Which types of interventions produce the greatest reductions in neighborhood-level early childhood obesity?</u>

We tested the hypothesis that neighborhoods that received multicomponent interventions would see greater declines in neighborhood-level prevalence. The results from the fixed-effects model of *change in obesity prevalence* on prior-year *intervention strategy count* can be seen in **Table 6.7**. For each additional type of intervention strategy neighborhoods received in the previous year, obesity prevalence would be expected to be 0.09% points (95% CI: 0.01, 0.16) lower compared to if those neighborhoods had not received any intervention, adjusting for prior-year intervention strategy count. For example, neighborhoods that received two different types of intervention strategies, such as government policies and group education, would be expected to have an obesity prevalence that was 0.18% percentage points lower than if those neighborhoods had not received any intervention, adjusting for prior-year intervention strategy count.

Table 6.7. Fixed-effects model of change in obesity prevalence on total types of intervention strategies neighborhoods received in the previous year

	Fixed-effects model, Change in obesity prevalence*^ β (95% CI)
Total types of intervention strategies$_{t-1}$	**-0.09 (-0.16, -0.01)**

*Refers to the difference between observed obesity prevalence with intervention and expected obesity prevalence without intervention
^Adjusted for prior-year intervention strategy count

A summary of the mean change in obesity prevalence by total types of intervention strategies neighborhoods received in the previous year can be seen in **Table 6.8**. Neighborhoods that received 1 intervention strategy type in the previous year had a mean obesity prevalence that was 0.43% points (95% CI: 0.14, 0.71) lower than if those neighborhoods had not received intervention. The more intervention strategy types neighborhoods received, the lower their mean obesity prevalence compared to if those neighborhoods had not received any intervention. Neighborhoods that received all 10 intervention strategy types would be expected to have a mean obesity prevalence that was 1.18% points (0.60, 1.77) lower than if those neighborhoods did not receive any intervention.

Table 6.8. Summary of mean change in obesity prevalence by total types of intervention strategies neighborhoods received in the previous year, 2006-2015

Total types of intervention strategies	N	Change in obesity prevalence*
		Mean (95% CI)
1_{t-1}	102	-0.43 (-0.71, -0.14)
2_{t-1}	209	-0.26 (-0.54, 0.03)
3_{t-1}	83	-0.30 (-0.59, 0.00)
4_{t-1}	100	-0.67 (-0.97, -0.37)
5_{t-1}	310	-0.62 (-0.79, -0.46)
6_{t-1}	283	-0.76 (-0.94, -0.57)
7_{t-1}	137	-0.79 (-1.03, -0.56)
8_{t-1}	100	-1.07 (-1.28, -0.86)
9_{t-1}	65	-1.10 (-1.34, -0.86)
10_{t-1}	22	-1.18 (-1.77, -0.60)

N= Number of ZIP Codes that received interventions between 2006 and 2015
*Refers to the difference between observed obesity prevalence with intervention and model-predicted obesity prevalence without intervention

We tested the hypothesis that neighborhoods that received multilevel interventions would see greater declines in neighborhood-level prevalence, that is, neighborhoods that received both macro-level and micro-level interventions. The results from the fixed-effects model of *change in obesity prevalence* on *level(s) of intervention* in the previous year can be seen in **Table 6.9**. Compared to neighborhoods that only received micro-level interventions in the previous year, neighborhoods that received macro-level intervention in the previous year would be expected to be have an obesity prevalence that was 2.49% points (95% CI: 1.36, 3.62) lower. Neighborhoods that received both macro-level and micro-level interventions in the previous year would be expected to be have an obesity prevalence that was 0.27% points (95% CI: 0.001, 0.54) lower compared to neighborhoods that only received micro-level interventions in the previous year.

Table 6.9. Fixed-effects model of change in obesity prevalence on level(s) of intervention that neighborhoods received in the previous year

Level(s) of intervention	Fixed-effects model, Change in obesity prevalence*^ β (95% CI)
Micro-level intervention only$_{t-1}$	Reference
Macro-level intervention only$_{t-1}$	**-2.49 (-3.62, -1.36)**
Both macro- and micro-level intervention$_{t-1}$	**-0.27 (-0.54, -0.001)**

*Refers to the difference between observed obesity prevalence with intervention and expected obesity prevalence without intervention
^Adjusted for prior-year intervention strategy count

A summary of the mean change in obesity prevalence by level(s) of intervention neighborhoods received in the previous year can be seen in **Table 6.10**. Neighborhoods that received any macro-level intervention in the preceding year had a mean obesity prevalence that was 1.36% points (95% CI: 0.36, 2.36) lower than if those neighborhoods had not received intervention. Neighborhoods that received any micro-level intervention in the previous year had a mean obesity prevalence that was 0.35% points (95% CI: 0.23, 0.47) lower than if those neighborhoods had not received intervention. Neighborhoods that received both macro- and micro-level interventions in the previous year had a mean obesity prevalence that was 0.75% points (95% CI: 0.85, 0.64) lower than if those neighborhoods had not received intervention.

Table 6.10. Summary of mean change in obesity prevalence by level(s) of intervention neighborhoods received in the previous year, 2006-2015

Level(s) of intervention	N	Change in obesity prevalence*
		Mean (95% CI)
Macro-level intervention(s) only$_{t-1}$	21	-1.36 (-2.36, -0.36)
Micro-level intervention(s) only$_{t-1}$	612	-0.35 (-0.47, -0.23)
Both macro-level and micro-level intervention(s)$_{t-1}$	970	-0.75 (-0.85, -0.64)

N= Number of ZIP Codes that received interventions between 2006 and 2015
*Refers to the difference between observed obesity prevalence with intervention and model-predicted obesity prevalence without intervention

We also examined whether macro-level or micro-level intervention strategies were more effective in reducing neighborhood-level obesity prevalence. The results from the fixed-effects model of *change in obesity prevalence* on the total *types of interventions neighborhoods received by level(s) of intervention* in the previous year can be seen in **Table 6.11**. For each additional macro-level intervention strategy *type* neighborhoods received in the previous year, obesity prevalence would be expected to be 0.26% points (95% CI: 0.09, 0.43) lower compared to if those neighborhoods had not received any intervention. For each additional micro-level intervention strategy type neighborhoods received in the previous year, obesity prevalence would be expected to be 0.05% points (95% CI: -0.02, 0.12) lower compared to if those neighborhoods had not received any intervention.

Table 6.11. Fixed-effects model of change in obesity prevalence on total types of interventions neighborhoods received by level(s) of intervention in the previous year

Total types of intervention by level(s) of intervention	Fixed-effects model, Change in obesity prevalence*^ β (95% CI)
Total types of macro-level intervention strategies used$_{t-1}$	**-0.26 (-0.43, -0.09)**
Total types of micro-level intervention strategies used$_{t-1}$	-0.05 (-0.12, 0.02)

*Refers to the difference between observed obesity prevalence with intervention and expected obesity prevalence without intervention
^Adjusted for prior-year intervention strategy count

A summary of the mean change in obesity prevalence by total types of macro-level intervention strategies and micro-level intervention strategies neighborhoods received in the previous year, respectively, can be seen in **Table 6.12** and **Table 6.13**. Neighborhoods that received 1 macro-level intervention strategy type in the previous year had a mean obesity prevalence that was 0.51% points (95% CI: 0.37, 0.66) lower than if those neighborhoods had not received intervention. The more macro-level intervention strategy types neighborhoods received in the previous year, the lower their mean obesity prevalence would be compared to if they had not received any intervention. Neighborhoods that received all 4 macro-level intervention strategy types in the previous year had a mean obesity prevalence that was 1.13% points (0.81, 1.45) lower than if those neighborhoods had not received any intervention. Neighborhoods that received 1 micro-level intervention strategy type in the previous year had a mean obesity prevalence that was 0.27% points (95% CI: 0.00, 0.53) lower than if those neighborhoods had not received intervention. The more micro-level intervention strategy types neighborhoods received in the previous year, the lower their mean obesity prevalence was compared to if they had not received any intervention. Neighborhoods that received all 6 micro-level intervention strategy types had a mean obesity prevalence that was 0.95% points (0.78, 1.13) lower than if those neighborhoods had not received any intervention.

Table 6.12. Summary of mean change in obesity prevalence by total types of macro-level intervention strategies neighborhoods received in the previous year

Total types of macro-level intervention strategies	N	Change in obesity prevalence* Mean (95% CI)
0_{t-1}	612	-0.35 (-0.47, -0.23)
1_{t-1}	590	-0.51 (-0.66, -0.37)
2_{t-1}	214	-1.14 (-1.35, -0.93)
3_{t-1}	131	-1.10 (-1.29, -0.91)
4_{t-1}	56	-1.13 (-1.45, -0.81)

N= Number of ZIP Codes that received interventions between 2006 and 2015
*Refers to the difference between observed obesity prevalence with intervention and model-predicted obesity prevalence without intervention

Table 6.13. Summary of mean change in obesity prevalence by total types of micro-level intervention strategies neighborhoods received in the previous year

Total types of micro-level intervention strategies	N	Change in obesity prevalence* Mean (95% CI)
0_{t-1}	213	-0.41 (-0.69, -0.12)
1_{t-1}	233	-0.27 (-0.53, 0.00)
2_{t-1}	170	-0.50 (-0.71, -0.28)
3_{t-1}	68	-1.00 (-1.36, -0.64)
4_{t-1}	158	-0.64 (-0.89, -0.38)
5_{t-1}	629	-0.70 (-0.81, -0.59)
6_{t-1}	132	-0.95 (-1.13, -0.78)

N= Number of ZIP Codes that received interventions between 2006 and 2015
*Refers to the difference between observed obesity prevalence with intervention and model-predicted obesity prevalence without intervention

Finally, we tested the hypothesis that neighborhoods that received interventions addressing more obesity-related behaviors (PA, diet, and BF) would see greater declines in neighborhood-level prevalence. The results from the fixed-effects model of *change in obesity* *prevalence* on obesity-related behaviors addressed by interventions in the previous year(s) can be seen in **Table 6.14**. Compared to neighborhoods that only received interventions addressing BF in the previous 3 years, obesity prevalence in neighborhoods that received interventions addressing PA, diet, PA and diet, PA and BF, or all 3 behaviors (PA, diet, and BF) was be expected to be lower. However, these effects were not found to be statistically significant. However, neighborhoods that received interventions addressing both BF and diet had an obesity prevalence that would be expected to be 0.61% points (95% CI: 0.11, 1.11) lower compared to if neighborhoods only received interventions addressing BF in the previous 3 years.

Table 6.14. Fixed-effects model of change in obesity prevalence on obesity-related behaviors addressed by interventions in the previous year(s)

Obesity-related behaviors addressed by interventions	Fixed-effects model, Change in obesity prevalence*^ β (95% CI)
Breastfeeding only$_{t-3}$	Reference
Physical Activity only$_{t-1}$	-0.74 (-1.72, 0.24)
Diet only$_{t-1}$	-0.60 (-1.42, 0.21)
Breastfeeding$_{t-3}$ & Physical Activity$_{t-1}$	-0.30 (-0.96, 0.38)
Breastfeeding$_{t-3}$ & Diet$_{t-1}$	**-0.61 (-1.11, -0.11)**
Physical Activity$_{t-1}$ & Diet$_{t-1}$	-0.10 (-0.58, 0.39)
Breastfeeding$_{t-3}$ & Physical Activity$_{t-1}$ & Diet$_{t-1}$	-0.32 (-0.66, 0.03)

*Refers to the difference between observed obesity prevalence with intervention and expected obesity prevalence without intervention
^Adjusted for prior-year intervention strategy count

A summary of mean change in obesity prevalence by obesity-related behaviors addressed by interventions in the previous year(s) can be seen in **Table 6.15**. Neighborhoods that received interventions targeting both diet in the previous year and BF in the previous 3 years saw the greatest decrease in obesity prevalence. Obesity prevalence in these neighborhoods was, on average, 1.02% points (95% CI: 0.76, 1.28) lower compared to if these neighborhoods had not received any intervention. Neighborhoods that received interventions only addressing PA saw a greater decrease in obesity prevalence compared to interventions only addressing diet or BF. Obesity prevalence in neighborhoods that only received interventions addressing PA was, on average, 0.96% points (95% CI: 0.43, 1.48) lower compared to if these neighborhoods had not received any intervention. For neighborhoods that only received interventions addressing diet, on average, obesity prevalence was 0.55% points (95% CI: 0.19, 0.91) lower compared to if these neighborhoods had not received any intervention. For neighborhoods that only received interventions addressing BF, on average, obesity prevalence was 0.71% points (95% CI: 0.48, 0.94) lower compared to if these neighborhoods had not received any intervention. Neighborhoods that received interventions addressing both PA and diet saw the smallest decrease in obesity prevalence, with obesity prevalence, on average, 0.37% points (95% CI: 0.19, 0.56) lower compared to if these neighborhoods had not received any intervention.

Table 6.15. Summary of mean change in obesity prevalence by obesity-related behaviors addressed by interventions in the previous year(s)

Obesity-related behaviors addressed by interventions	N	Change in obesity prevalence* Mean (95% CI)
Breastfeeding only$_{t-3}$	84	-0.71 (-0.94, -0.48)
Physical Activity only$_{t-1}$	46	-0.96 (-1.48, -0.43)
Diet only$_{t-1}$	87	-1.05 (-1.54, -0.57)
Breastfeeding$_{t-3}$ × Physical Activity$_{t-1}$	28	-0.74 (-1.27, -0.20)
Breastfeeding$_{t-3}$ × Diet$_{t-1}$	146	-1.02 (-1.28, -0.76)
Physical Activity$_{t-1}$ × Diet$_{t-1}$	327	-0.45 (-0.65, -0.25)
Breastfeeding$_{t-3}$ × Physical Activity$_{t-1}$ × Diet$_{t-1}$	392	-0.77 (-0.88, -0.66)

N= Number of ZIP Codes that received interventions between 2006 and 2015
*Refers to the difference between observed obesity prevalence with intervention and model-predicted obesity prevalence without intervention

Discussion

Using machine learning techniques, we were able to create a reasonable hypothetical comparison group of neighborhoods that did not receive obesity-related interventions. The trained model's predictions of obesity prevalence for 6 years of data (that were not included in the original model development), were highly correlated with the observed data, in both the ZIP Codes included in the training model (correlation 0.80) and those excluded from the development for use as a completely separated validation sample (correlation 0.84), demonstrating that the model had validity to make good predictions, based on sociodemographic data and prior-year obesity prevalence.

We found evidence demonstrating that place-based interventions do contribute to declines in population-level rates of childhood obesity. The higher the intervention strategy count neighborhoods received, the greater the decline in neighborhood-level early childhood obesity prevalence. For each additional intervention strategy neighborhoods received in the previous year, obesity prevalence would decrease by 0.04% points (95% CI: 0.03, 0.06). In CHAPTER 4, we showed that in 2013-2015, neighborhoods in LAC received, on average, about 50 intervention strategies or more. For a neighborhood that received 50 intervention strategies in the previous year would, for example, obesity prevalence would decrease by about 1.5%. Our finding is consistent with the findings from the Healthy Communities Study, in which greater intensity of community-based programs and policies, characterized by behavior change strategy, duration, reach and other dimensions, were related to lower childhood BMI.[266]

Multicomponent and multilevel interventions

Leading health organizations including the National Academy of Medicine (formerly the Institute of Medicine) and the World Health Organization recommend comprehensive and multi-level approaches for obesity prevention, including programs and policies at the community level.[267] Models like the socioecological model[14] offer a framework for better understanding the multiple and interacting activities that determine health behaviors, and are being used more often to implement comprehensive approaches to addressing childhood obesity. Recent reviews suggest comprehensive programs and policies at multiple levels may be more promising for reducing obesity in children than single-level interventions that take less comprehensive approaches.[21-23,25,268] However, the evidence, while increasing, is still limited and little evidence is available on the extent to which multicomponent and multilevel interventions improve obesity outcomes, as all of the published reviews are based on 10 or fewer published community-level studies.

Using a rich intervention database that characterized all major obesity-related initiatives implemented in LAC over 10 years, neighborhood-level prevalence estimates of all WIC-enrolled preschool-aged children in LAC, and the application of sophisticated machine modelling techniques to create a hypothetical "comparison" group, we found that multicomponent interventions, that is, interventions that used more intervention strategy types, were effective at reducing population-level prevalence of early childhood obesity, and this relationship followed a dose response. For each additional intervention strategy type a neighborhood received, neighborhood-level prevalence of early childhood obesity would decrease by 0.08% points (95% CI: 0.03, 0.13). From 2006-2015, there were over 600 instances where a ZIP Code received more than 5 intervention strategies (in ZIP Code-years). For a neighborhood that received at least 5 intervention strategy types in the previous year, early childhood obesity prevalence would decrease by at least 0.4%. Our findings support the recommendation for multicomponent interventions.

Based on our understanding of the socioecological framework,[14] we expect multilevel interventions to decrease prevalence of obesity, more so than single-level interventions yet our study findings did not support this. Based on our study findings, macro-level interventions played an important role in contributing to neighborhood-level declines in early childhood obesity prevalence. Neighborhoods that received any macro-level intervention saw obesity prevalence decrease by 1.03% points (95% CI: 0.30, 1.76). In our analysis, the effect of receiving any micro-level intervention in decreasing neighborhood-level obesity prevalence was inconclusive. Contrary to our hypothesis, we found that for neighborhoods that received multilevel interventions (both macro- and micro-level) saw a decrease in obesity prevalence of

0.18% points compared to 1.03% points for neighborhoods that only received macro-level interventions. However, we would like to point out that in our analysis, there were only 21 observations (in ZIP Code-years) where neighborhoods only received macro-level interventions, whereas 970 observations consisted of neighborhoods that received both macro-level and micro-level interventions. Nonetheless, our analysis supports the larger effect of macro-level interventions on reducing neighborhood-level prevalence of early childhood obesity, and we found, for each additional macro-level intervention strategy type a neighborhood received, obesity prevalence decreased by 0.30% points (95% CI: 0.18, 0.42) compared to a decrease of 0.06% points (95% CI: 0.01, 0.11) for each additional micro-level intervention strategy type a neighborhood received. A neighborhood receiving all 4 macro-level intervention strategy types could potentially have its obesity prevalence reduced by about 1.2%. Comparatively, a neighborhood that received all 6 micro-level strategies could potentially have its obesity prevalence reduced by about 0.36%.

Micro-level intervention strategies that simply involve providing parents with information, guidance, or encouragement, relies on parents to be able and motivated to engage in behavior change and adopt healthier lifestyles for themselves and their children. However, parents in more under-resourced settings may face significant barriers, such as lower health literacy, and so may engage less with public health information. Furthermore, according to the theory of fundamental causes,[77] affluent parents are more likely to have the material resources to be able to afford more expensive, healthier foods, and the time resources to source and prepare them.[269] There is also evidence suggesting that mothers of young children with higher educational attainment are better equipped to navigate a poor food environment, thereby resulting diets that are relatively more healthy given the circumstances.[270] Adams et al. describe such findings using the concept of ‘agency’ – the personal resources individuals have to use in order to benefit from interventions, such as cognitive, psychological and material resources and time.[271] Macro-level interventions that require recipients to use little or no agency to improve dietary habits by, for example, changing the physical environment or food provision may be more effective and equitable than those that require high use of agency. Success with interventions that focus on changing parents’ behavior, which in turn influences the behaviors of their child, is more difficult without an environment that supports those changes and is conducive to healthy choices. Although we did not find that neighborhoods receiving multilevel interventions saw greater declines in obesity prevalence, we did find that both macro- and micro-level interventions were successful in decreasing neighborhood-level obesity prevalence, and support the recommendation for multilevel interventions.

Interventions addressing multiple obesity-related behaviors

Levels of PA, diet, and being breastfed are all important behaviors that influence the development of early childhood obesity. Neighborhoods that received any intervention addressing PA, diet, and BF saw obesity prevalence decrease by 0.55% points (95% CI: 0.04, 1.06), 0.58% points (95% CI: 0.19, 0.98), and 0.21% points (95% CI: -0.19, 0.61), respectively. Interventions that address diet may be slightly more effective than interventions that address PA. Based on our findings, the effects of interventions addressing BF are inconclusive. Despite our hypothesis, we did not find evidence that neighborhoods that received interventions addressing more obesity-related behaviors (PA, diet, and BF) saw greater declines in neighborhood-level prevalence. The estimate of interaction term between interventions addressing all 3 obesity-related behaviors was in the hypothesized direction (-0.78% points, suggesting a decrease in obesity prevalence), but was not statistically significant.

In the published literature, the effects of interventions that combine both diet and PA have been mixed. A Cochrane review of 16 randomized control trials (RCT) found that interventions that addressed both diet and PA among children 0 to 5 years reduced BMI and standardized BMI (zBMI) compared with control, but interventions addressing either diet or PA alone did not.[11] Based on 14 RCTs, PA interventions reduced BMI but had little or no effect on zBMI for children aged 6 to 12 years compared with controls, whereas interventions addressing both diet and PA could potentially reduce zBMI, but diet interventions had little impact on zBMI or BMI.[11] Another Cochrane review, based on 44 RCTs among adolescents 12 to 17 years, found low quality evidence suggesting that interventions addressing both diet and PA could potentially reduce measures of BMI, and moderate quality evidence that they reduced weight in overweight or obese adolescents when compared with no treatment or controls.[12] A systematic review and meta-analysis reported that the highest proportion of significant and favorable impacts on adiposity-related outcomes was attributable to diet-only interventions, while the lowest proportion of successes was attributable to PA-only interventions.[25] The effectiveness of interventions addressing breastfeeding on reductions in childhood obesity have also been mixed, with reviews either finding inconclusive evidence or no evidence to support their effects.[272-274]

Study strengths and limitations

This study had a number of strengths. Access to an intervention database that classified programs and projects facilitated the concept of *intervention strategy count* which allowed us to quantify exposure of a community to various intervention strategies implemented simultaneously by various obesity-related initiatives in LAC. Use of this intervention database allowed us to examine the cumulative effects of simultaneous interventions on declines in population-level rates of childhood obesity. The published literature reviews on evaluation of place-based interventions relied on published studies, and often only included RCTs. Authors of these reviews have noted that studies on macro-level interventions are largely missing in the published literature.[272,275] The application of sophisticated modelling techniques to a unique intervention database provided us with an opportunity to establish the role of place-based interventions in contributing to declines in population-level rates of childhood obesity, which has not previously been shown in the published literature. Finally, measured heights and weights of high validity[149] were used to calculate obesity prevalence estimates.

This study had several limitations. First, our machine learning model's predictions were highly correlated with the observed data (correlation of 0.80-0.84). However, since our predictions were not entirely accurate, we acknowledge error in the model-predicted obesity prevalence estimates under no intervention. Since all of our findings rest of the outcome *change in obesity prevalence*, which relies on these model-predicted obesity prevalence estimates, we must be cognizant of this as we interpret the study findings. Notably, our intervention database is not comprehensive of all obesity-related interventions that took place in LAC. It focused on interventions implemented by major funders and health organizations tackling obesity in LAC, and part of the intervention data collection focused on WIC clinics in 8 regions of LAC where the majority of WIC families reside (**Appendix 1**). While the intervention strategy count allows for the quantification of exposure to various strategies, it does not consider the reach of a strategy. Third, intervention data were available only at the ZIP Code level, necessitating neighborhoods to be defined by ZIP Codes; these are relatively large geographic spaces and consequentially more likely to display heterogeneous neighborhood effects.[192] Sociodemographic data obtained from the US Census Bureau are also only available at the level of ZCTAs, not ZIP Codes.

Conclusion

This study illustrates the novel application of machine learning techniques to public health and evaluation research. Machine learning can be used to create a reasonable hypothetical comparison group of neighborhoods, permitting us to begin to evaluate and understand the complexities of the community-level, place-based intervention landscape. Our unique analytical approach allows us to mimic a natural experiment, which refers to "naturally occurring circumstances in which different populations may or may not be exposed to a potentially causal factor or intervention such that the circumstances resemble a true experiment in which participants may be assigned to exposed or unexposed groups".[276] Natural experiments hold advantages for external validity because they reflect the real-world challenges of implementing programs and policies that cannot be assigned in the unusual circumstances of community trials or effectiveness studies.[276] The promising results from this study suggest that place-based interventions are likely to be responsible for reductions in neighborhood-level early childhood obesity for low-income, WIC-enrolled children. Our findings provide a clear direction towards better understanding what types of interventions can impact population-level rates of early childhood obesity.

CHAPTER 7: Conclusion

Over the last decade, rates of obesity among children 2 to 5 years in the United States (US), at certain points in time, seemed to be heading in the right direction. Between 2009-2010 and 2011-2012, obesity prevalence significantly decreased from 12.1% to 8.4%.[1] However, obesity prevalence increased to 9.4% in 2013-2014, and continued to increase again, reaching 13.9% in 2015-2016, the latest estimate available.[1,3] Obesity prevalence among 2- to 5-year-old children is the highest it has been since 2003-2004.[1]

For WIC-enrolled children 2 to 5 years, the trends in obesity look more positive. From a high of 15.9% in 2010, obesity rates have continued to decrease yearly, albeit in small increments.[50,51] For the first time in the last 2 decades, obesity prevalence among WIC-enrolled children was not higher than it was for the overall US population of children 2 to 5 years. In 2016, the most recent estimate available, obesity prevalence reached 13.9% for both the WIC-enrolled and the overall US population of preschool-aged children.[3,50,51]

Though the trends for WIC-enrolled children in LAC follow the trends seen for all WIC-enrolled children in the US, obesity prevalence remains significantly higher for preschool-aged children in Los Angeles County (LAC). Obesity prevalence was 20.2% in 2010, compared to 15.9% for all WIC-enrolled children in the US.[50,147] Since 2010, obesity rates in LAC have continued to steadily decrease, reaching 18% in 2016, although it remains 4% higher than both the WIC-enrolled and the overall US population of preschool-aged children.[3,50,51,147]

With all of the money, resources, and carefully planned place-based interventions that have been implemented to address obesity,[18-25] these trends suggest that these efforts may be paying off, as many of these place-based interventions have been targeting under-resourced neighborhoods.[22,139] In CHAPTER 4, we found that in LAC, resources for obesity control were more heavily directed towards poorer neighborhoods, and neighborhoods with predominantly Black or Hispanic residents. These findings align with Frohlich and Potvin's approach to public health interventions. Frohlich and Potvin's "vulnerable population" approach underscores the importance of interventions targeting disadvantaged populations, rather than the entire population, to reduce health disparities.[130] In LAC, place-based interventions addressing obesity may very well have contributed to the trends in obesity prevalence we have seen among WIC-enrolled preschool-aged children. However, large disparities in neighborhood-level prevalence of early childhood obesity remain. In CHAPTER 4, we showed that in 2015, the neighborhood with the highest prevalence of early childhood obesity (30.51%) had a prevalence of obesity that was more than six times greater than the neighborhood with the lowest prevalence (4.43%). To reduce these disparities, neighborhoods most burdened by early childhood obesity should be prioritized for obesity control.

However, current data limitations, specifically, the lag times in the availability of surveillance data, make it very difficult for policymakers and public health practitioners to identify neighborhoods most burdened by early childhood obesity. In CHAPTER 5, we illustrated a novel approach to identify these neighborhoods in a timely manner. Machine learning is a relatively unfamiliar approach within public health, but we have attempted to present it as a very useful tool, with many applications. In CHAPTER 5, we found that, at a minimum, the last obesity prevalence estimates available (due to lag times in data availability), could reasonably predict future prevalence of obesity at the ZIP Code level. For example, to estimate a neighborhood's obesity prevalence in 2020, we could use the latest obesity prevalence estimate available, that is, that neighborhood's prevalence of obesity in 2017. However, using machine learning techniques, and including publicly available sociodemographic data to our model, we can improve the accuracy of making future projections of neighborhood-level prevalence of obesity. This simple application of machine learning algorithms can help policymakers quickly identify those neighborhoods most burdened by obesity— where obesity prevention efforts are most needed. The predictive capabilities of machine learning algorithms can be applied to many different types of scenarios, and can prove to be useful in our public health "toolbox".

In CHAPTER 6, we illustrated another application of machine learning. We applied machine learning techniques to create a hypothetical comparison group to allow us to evaluate place-based interventions addressing obesity. As we have discussed throughout this dissertation, the socioecological model[14] offers a framework for better understanding the multiple and interacting factors that shape obesity-related behaviors. As public health continues to adopt ecological frameworks to address obesity and implement comprehensive approaches at multiple levels of the socioecological model— while grounded in theory and incredibly important— these types of interventions make the evaluation exceedingly difficult. With this type of multilevel, multicomponent, longer-term interventions, it becomes difficult to keep track of the specific components of the intervention that are being implemented, when they are being implemented, where, with what intensity, the population it is reaching, etc. This underscores the importance of an intervention database like the one created for ECOSyS. Given the complexities of the intervention landscape in the *real world* and the difficulties in evaluating such complex interventions provides greater justification as to why, as public health researchers, we should be open to novel methods and frameworks, borrowing from other disciplines, to creatively overcome the challenges that are involved when trying to do this kind of evaluation research.

We have illustrated how machine learning techniques can be applied to public health and evaluation research. Application of machine learning to predict future prevalence of diseases or risk factors, not just obesity, can support policymakers and public health practitioners identify communities in greatest need of intervention. Our overall findings suggest that place-based interventions to address early childhood obesity have been working in LAC, which is not only an encouraging finding, but underscores that importance of continuing to do this work and work towards reducing disparities in childhood obesity, as well as reducing overall rates or early childhood obesity.

REFERENCES

1. Ogden CL, Carroll MD, Lawman HG, et al. Trends in obesity prevalence among children and adolescents in the United States, 1988-1994 through 2013-2014. *JAMA.* 2016;315(21):2292-2299.
2. Organization for Economic Co-operation and Development. Obesity Update 2017. http://www.oecd.org/health/obesity-update.htm. Published 2017. Accessed October 20, 2020.
3. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity among adults and youth: United States, 2015-2016. *NCHS Data Brief.* 2017(288):1-8.
4. World Health Organization. Noncommunicable diseases: Childhood overweight and obesity. https://www.who.int/news-room/q-a-detail/noncommunicable-diseases-childhood-overweight-and-obesity. Published 2020. Accessed October 19, 2020.
5. Inchley J, Currie D, Young T, et al. *Growing up unequal: Gender and socioeconomic differences in young people's health and well-being, Health Behaviour in School-aged Children (HBSC) Study: International report from the 2013/2014 survey.* Copenhagen: WHO Regional Office for Europe;2016.
6. Ludwig DS. Childhood obesity — The shape of things to come. *N Engl J Med.* 2007;357(23):2325-2327.
7. Singh AS, Mulder C, Twisk JW, van Mechelen W, Chinapaw MJ. Tracking of childhood overweight into adulthood: A systematic review of the literature. *Obes Rev.* 2008;9(5):474-488.
8. Simmonds M, Llewellyn A, Owen CG, Woolacott N. Predicting adult obesity from childhood obesity: A systematic review and meta-analysis. *Obes Rev.* 2016;17(2):95-107.
9. Kit BK, Ogden CL, Flegal KM. Epidemiology of Obesity. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology.* New York, NY: Springer New York; 2014:2229-2262.
10. Hu F. *Obesity Epidemiology.* New York: Oxford University Press; 2008.
11. Brown T, Moore THM, Hooper L, et al. Interventions for preventing obesity in children. *Cochrane Database Syst Rev.* 2019(7).
12. Al-Khudairy L, Loveman E, Colquitt JL, et al. Diet, physical activity and behavioural interventions for the treatment of overweight or obese adolescents aged 12 to 17 years. *Cochrane Database Syst Rev.* 2017(6).
13. Mead E, Brown T, Rees K, et al. Diet, physical activity and behavioural interventions for the treatment of overweight or obese children from the age of 6 to 11 years. *Cochrane Database Syst Rev.* 2017(6).
14. Bronfenbrenner U. *The Ecology of Human Development: Experiments by Nature and Design.* Cambridge, MA: Harvard University Press; 1979.
15. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. *Health Educ Q.* 1988;15(4):351-377.
16. Stokols D. Translating social ecological theory into guidelines for community health promotion. *Am J Health Promot.* 1996;10(4):282-298.
17. Swinburn B, Egger G, Raza F. Dissecting obesogenic environments: The development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Prev Med.* 1999;29(6 Pt 1):563-570.

18. The California Endowment. Building Healthy Communities. www.calendow.org/building-healthy-communities/. Accessed October 19, 2020.
19. Centers for Disease Control and Prevention. Community Transformation Grants (2011–2014). https://www.cdc.gov/nccdphp/dch/programs/communitytransformation/. Updated March 7, 2017. Accessed October 19, 2020.
20. Robert Wood Johnson Foundation. Robert Wood Johnson Foundation Doubles Its Commitment to Helping All Children Grow Up at a Healthy Weigh. https://www.rwjf.org/en/library/articles-and-news/2015/02/rwjf_doubles_commitment_to_healthy_weight_for_children.html. Published 2015. Accessed October 19, 2020.
21. Wolfenden L, Wyse R, Nichols M, Allender S, Millar L, McElduff P. A systematic review and meta-analysis of whole of community interventions to prevent excessive population weight gain. *Prev Med.* 2014;62:193-200.
22. Boelsen-Robinson T, Peeters A, Beauchamp A, Chung A, Gearon E, Backholer K. A systematic review of the effectiveness of whole-of-community interventions by socioeconomic position. *Obes Rev.* 2015;16(9):806-816.
23. Ewart-Pierce E, Mejía Ruiz MJ, Gittelsohn J. "Whole-of-community" obesity prevention: A review of challenges and opportunities in multilevel, multicomponent interventions. *Curr Obes Rep.* 2016;5(3):361-374.
24. Scott-Sheldon LAJ, Hedges LV, Cyr C, et al. Childhood Obesity Evidence Base Project: A systematic review and meta-analysis of a new taxonomy of intervention components to improve weight status in children 2–5 years of age, 2005–2019. *Child Obes.* 2020;16(S2):S2-21-S22-48.
25. Wang Y, Cai L, Wu Y, et al. What childhood obesity prevention programmes work? A systematic review and meta-analysis. *Obes Rev.* 2015;16(7):547-565.
26. Economos CD, Hyatt RR, Must A, et al. Shape Up Somerville two-year results: A community-based environmental change intervention sustains weight reduction in children. *Prev Med.* 2013;57(4):322-327.
27. Han JC, Lawlor DA, Kimm SY. Childhood obesity – 2010: Progress and challenges. *Lancet.* 2010;375(9727):1737-1748.
28. Daniels SR. Complications of obesity in children and adolescents. *Int J Obes (Lond).* 2009;33 Suppl 1:S60-65.
29. Koyuncuoğlu Güngör N. Overweight and obesity in children and adolescents. *J Clin Res Pediatr Endocrinol.* 2014;6(3):129-143.
30. Klish WJ, Skelton JA. Comorbidities and complications of obesity in children and adolescents. https://www.uptodate.com/contents/comorbidities-and-complications-of-obesity-in-children-and-adolescents. Updated Mar 31, 2020. Accessed October 30, 2020.
31. Cohen AK, Rai M, Rehkopf DH, Abrams B. Educational attainment and obesity: A systematic review. *Obes Rev.* 2013;14(12):989-1005.
32. World Health Organization. Obesity and overweight: Key facts. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight. Published 2020. Updated April 1, 2020. Accessed October 30, 2020.
33. Finkelstein EA, Graham WCK, Malhotra R. Lifetime direct medical costs of childhood obesity. *Pediatrics.* 2014;133(5):854-862.

34. Wells JCK. *The Evolutionary Biology of Human Body Fatness: Thrift and Control.* Cambridge: Cambridge University Press; 2009.
35. Simmonds M, Burch J, Llewellyn A, et al. The use of measures of obesity in childhood for predicting obesity and the development of obesity-related diseases in adulthood: A systematic review and meta-analysis. *Health Technol Assess.* 2015;19(43):1-336.
36. Freedman DS, Khan LK, Serdula MK, Dietz WH, Srinivasan SR, Berenson GS. Inter-relationships among childhood BMI, childhood height, and adult obesity: The Bogalusa Heart Study. *Int J Obes Relat Metab Disord.* 2004;28(1):10-16.
37. Berrington de Gonzalez A, Hartge P, Cerhan JR, et al. Body-Mass Index and mortality among 1.46 million white adults. *N Engl J Med.* 2010;363(23):2211-2219.
38. NHLBI Obesity Education Initiative Expert Panel on the Identification Evaluation and Treatment of Obesity in Adults (US). *Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: The evidence report.* Vol 98-4083. Bethesda, MD: National Heart, Lung, and Blood Institute; 1998.
39. McGee DL. Body Mass Index and mortality: A meta-analysis based on person-level data from twenty-six observational studies. *Ann Epidemiol.* 2005;15(2):87-97.
40. Kopelman P. Health risks associated with overweight and obesity. *Obes Rev.* 2007;8:13-17.
41. Livingstone B. Epidemiology of childhood obesity in Europe. *Eur J Pediatr.* 2000;159 Suppl 1:S14-34.
42. Kipping RR, Jago R, Lawlor DA. Obesity in children. Part 1: Epidemiology, measurement, risk factors, and screening. *BMJ.* 2008;337:a1824.
43. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ.* 2007;85(9):660-667.
44. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: International survey. *BMJ.* 2000;320(7244):1240-1243.
45. Kuczmarski RJ, Ogden CL, Grummer-Strawn LM, et al. CDC growth charts: United States. *Adv Data.* 2000(314):1-27.
46. Kuczmarski RJ, Ogden CL, Guo SS, et al. 2000 CDC growth charts for the United States: Methods and development. *Vital Health Stat 11.* 2002(246):1-190.
47. Hawkins SS, Oken E, Gillman MW. Early in the Life Course: Time for Obesity Prevention. In: Halfon N, Forrest CB, Lerner RM, Faustman EM, eds. *Handbook of Life Course Health Development.* Cham (CH): Springer; 2018:169-196.
48. Institute of Medicine. *Preventing Childhood Obesity: Health in the Balance.* Washington, DC: The National Academies Press; 2005.
49. Ogden CL, Lamb MM, Carroll MD, Flegal KM. Obesity and socioeconomic status in children and adolescents: United States, 2005-2008. *NCHS Data Brief.* 2010(51):1-8.
50. Pan L, Freedman DS, Sharma AJ, et al. Trends in obesity among participants aged 2-4 years in the Special Supplemental Nutrition Program for Women, Infants, and Children United States, 2000-2014. *Morb Mortal Wkly Rep.* 2016;65(45):1256-1260.

51. Pan L, Freedman DS, Park S, Galuska DA, Potter A, Blanck HM. Changes in obesity among US children aged 2 through 4 years enrolled in WIC during 2010-2016. *JAMA.* 2019;321(23):2364-2366.
52. U.S. Census Bureau. QuickFacts: Los Angeles County, California; UNITED STATES. https://www.census.gov/quickfacts/fact/table/losangelescountycalifornia,US/PST045219. Accessed October 19, 2020.
53. Bandy M. The WIC program in California. Public Policy Institute of California. https://www.ppic.org/content/pubs/jtf/JTF_WICJTF.pdf. Published 2015. Accessed November 1, 2020.
54. LA County WIC Data. Los Angeles County WIC Data. http://lawicdata.org/about/. Accessed October 31, 2020.
55. National Collaborative on Childhood Obesity Research. Catalogue of surveillance systems. https://tools.nccor.org/css/. Accessed October 31, 2020.
56. Centers for Disease Control and Prevention. Overweight & obesity: Obesity among WIC-enrolled young children. https://www.cdc.gov/obesity/data/obesity-among-WIC-enrolled-young-children.html. Updated November 22, 2019. Accessed October 30, 2020.
57. LA County WIC Data. Childhood obesity. https://lawicdata.org/data-research/topics/obesity-2/. Accessed October 31, 2020.
58. Bouchard C, Perusse L, Rice T, Rao DC. Genetics of Human Obesity. In: Bray GA, Bouchard C, eds. *Handbook of Obesity: Etiology and Pathophysiology.* Second ed. New York: Marcel Dekker; 2004.
59. Maes HH, Neale MC, Eaves LJ. Genetic and environmental factors in relative body weight and human adiposity. *Behav Genet.* 1997;27(4):325-351.
60. Jaacks LM, Vandevijvere S, Pan A, et al. The obesity transition: Stages of the global epidemic. *Lancet Diabetes Endocrinol.* 2019;7(3):231-240.
61. Fryar CD, Carroll MD, Ogden CL, Division of Health and Nutrition Examination Surveys. Prevalence of overweight, obesity, and extreme obesity among adults aged 20 and over: United States, 1960-1962 through 2013-2014. NCHS Health E-Stats Web site. https://www.cdc.gov/nchs/data/hestat/obesity_adult_13_14/obesity_adult_13_14.pdf. Published 2016. Accessed October 30, 2020.
62. Braveman P, Gottlieb L. The social determinants of health: It's time to consider the causes of the causes. *Public Health Rep.* 2014;129(Suppl 2):19-31.
63. Birch L, Savage JS, Ventura A. Influences on the development of children's eating behaviours: From infancy to adolescence. *Can J Diet Pract Res.* 2007;68(1):s1-s56.
64. Hill AJ. Developmental issues in attitudes to food and diet. *Proc Nutr Soc.* 2002;61(2):259-266.
65. Lindsay AC, Sussner KM, Kim J, Gortmaker S. The role of parents in preventing childhood obesity. *Future Child.* 2006;16(1):169-186.
66. Faith MS, Scanlon KS, Birch LL, Francis LA, Sherry B. Parent-child feeding strategies and their relationships to child eating and weight status. *Obes Res.* 2004;12(11):1711-1722.
67. Hughes SO, Power TG, Orlet Fisher J, Mueller S, Nicklas TA. Revisiting a neglected construct: Parenting styles in a child-feeding context. *Appetite.* 2005;44(1):83-92.

68. Patrick H, Nicklas TA. A review of family and social determinants of children's eating patterns and diet quality. *J Am Coll Nutr.* 2005;24(2):83-92.
69. Committee on Physical Activity and Physical Education in the School Environment, Food and Nutrition Board, Institute of Medicine. *Educating the Student Body: Taking Physical Activity and Physical Education to School.* Washington, DC: National Academies Press (US); 2013.
70. Petersen TL, Møller LB, Brønd JC, Jepsen R, Grøntved A. Association between parent and child physical activity: A systematic review. *Int J Behav Nutr Phys Act.* 2020;17(1):67.
71. Hutchens A, Lee RE. Parenting practices and children's physical activity: An integrative review. *J Sch Nurs.* 2018;34(1):68-85.
72. Sebire SJ, Jago R, Wood L, Thompson JL, Zahra J, Lawlor DA. Examining a conceptual model of parental nurturance, parenting practices and physical activity among 5–6 year olds. *Soc Sci Med.* 2016;148:18-24.
73. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice Hall; 1986.
74. Pearson N, Biddle SJ, Gorely T. Family correlates of fruit and vegetable consumption in children and adolescents: A systematic review. *Public Health Nutr.* 2009;12(2):267-283.
75. Cockerham WC. Health lifestyle theory and the convergence of agency and structure. *J Health Soc Behav.* 2005;46(1):51-67.
76. Adler NE, Boyce T, Chesney MA, et al. Socioeconomic status and health. The challenge of the gradient. *Am Psychol.* 1994;49(1):15-24.
77. Phelan JC, Link BG, Tehranifar P. Social conditions as fundamental causes of health inequalities: Theory, evidence, and policy implications. *J Health Soc Behav.* 2010;51(S):S28-S40.
78. Feagin JR. *Racist America: Roots, current realities, and future reparations.* 3 ed. New York and London: Routledge; 2014.
79. Phelan JC, Link BG. Is racism a fundamental cause of inequalities in health? *Annu Rev Sociol.* 2015;41(1).
80. Hill JO, Wyatt HR, Reed GW, Peters JC. Obesity and the environment: Where do we go from here? *Science.* 2003;299(5608):853-855.
81. Voigt K, Nicholls SG, Williams G. Childhood Obesity and the "Obesogenic Environment". In: *Childhood Obesity: Ethical and Policy Issues.* New York, NY: Oxford University Press; 2014.
82. Papas MA, Alberg AJ, Ewing R, Helzlsouer KJ, Gary TL, Klassen AC. The built environment and obesity. *Epidemiol Rev.* 2007;29(1):129-143.
83. Sallis JF, Nader PR, Rupp JW, Atkins CJ, Wilson WC. San Diego surveyed for heart-healthy foods and exercise facilities. *Public Health Rep.* 1986;101(2):216-219.
84. Morland K, Diez Roux AV, Wing S. Supermarkets, other food stores, and obesity: The Atherosclerosis Risk in Communities Study. *Am J Prev Med.* 2006;30(4):333-339.
85. Prentice AM, Jebb SA. Fast foods, energy density and obesity: A possible mechanistic link. *Obes Rev.* 2003;4(4):187-194.
86. Lin B, Frazão E, Guthrie J. *Away-from-home food increasingly important to quality of American diet.* Washington, DC: U.S. Department of Agriculture; 1999.

87. Bowman SA, Vinyard BT. Fast food consumption of U.S. adults: Impact on energy and nutrient intakes and overweight status. *J Am Coll Nutr.* 2004;23(2):163-168.
88. Glanz K, Basil M, Maibach E, Goldberg J, Snyder DAN. Why Americans eat what they do: Taste, nutrition, cost, convenience, and weight control concerns as influences on food consumption. *J Am Diet Assoc.* 1998;98(10):1118-1126.
89. Townshend T, Lake A. Obesogenic environments: Current evidence of the built and food environments. *Perspect Public Health.* 2017;137(1):38-44.
90. Huybrechts I, De Bourdeaudhuij I, De Henauw S. Environmental Factors: Opportunities and Barriers for Physical Activity, and Healthy Eating. In: Moreno LA, Pigeot I, Ahrens W, eds. *Epidemiology of Obesity in Children and Adolescents: Prevalence and Etiology.* New York: Springer; 2011.
91. Bennett GG, Wolin KY, Duncan DT. Social Determinants of Obesity. In: Hu F, ed. *Obesity Epidemiology.* New York: Oxford University Press; 2008.
92. Feng J, Glass TA, Curriero FC, Stewart WF, Schwartz BS. The built environment and obesity: A systematic review of the epidemiologic evidence. *Health Place.* 2010;16(2):175-190.
93. Adams MA, Ding D, Sallis JF, et al. Patterns of neighborhood environment attributes related to physical activity across 11 countries: A latent class analysis. *Int J Behav Nutr Phys Act.* 2013;10:34.
94. Farley TA, Meriwether RA, Baker ET, Rice JC, Webber LS. Where do the children play? The influence of playground equipment on physical activity of children in free play. *J Phys Act Health.* 2008;5(2):319-331.
95. Huang J-H, Hipp JA, Marquet O, et al. Neighborhood characteristics associated with park use and park-based physical activity among children in low-income diverse neighborhoods in New York City. *Prev Med.* 2020;131:105948.
96. Reimers AK, Schoeppe S, Demetriou Y, Knapp G. Physical activity and outdoor play of children in public playgrounds— Do gender and social environment matter? *Int J Environ Res Public Health.* 2018;15(7):1356.
97. Kawachi I. Social capital and community effects on population and individual health. *Ann N Y Acad Sci.* 1999;896(1):120-130.
98. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science.* 1997;277(5328):918-924.
99. Kawachi I, Berkman LF. Social Cohesion, Social Capital, and Health. In: Berkman LF, Kawachi I, eds. *Social Epidemiology.* New York, NY: Oxford University Press; 2000.
100. Sampson RJ, Morenoff JD, Earls F. Beyond social capital: Spatial dynamics of collective efficacy for children. *Am Sociol Rev.* 1999;64(5):633-660.
101. Hechter M, Opp K. *Social Norms.* New York, NY: Russell Sage; 2001.
102. Carter MA, Dubois L. Neighbourhoods and child adiposity: A critical appraisal of the literature. *Health Place.* 2010;16(3):616-628.
103. Umberson D, Crosnoe R, Reczek C. Social relationships and health behavior across life course. *Annu Rev Sociol.* 2010;36:139-157.
104. Pebley AR, Sastry N. Neighborhoods, Poverty, and Children's Well-being. In: Grusky DB, Ku MC, Szelenyi S, eds. *Social Stratification: Class, Race, and Gender in*

Sociological Perspective. Third ed. New York, NY: Russell Sage Foundation; 2008:360-370.
105. Sallis JF, Slymen DJ, Conway TL, et al. Income disparities in perceived neighborhood built and social environment attributes. *Health Place.* 2011;17(6):1274-1283.
106. Frank L, Kerr J, Saelens B, Sallis J, Glanz K, Chapman J. Food outlet visits, physical activity and body weight: Variations by gender and race-ethnicity. *Br J Sports Med.* 2009;43(2):124.
107. Lovasi GS, Hutson MA, Guerra M, Neckerman KM. Built environments and obesity in disadvantaged populations. *Epidemiol Rev.* 2009;31(1):7-20.
108. Bower KM, Thorpe RJ, Rohde C, Gaskin DJ. The intersection of neighborhood racial segregation, poverty, and urbanicity and its impact on food store availability in the United States. *Prev Med.* 2014;58:33-39.
109. James P, Arcaya MC, Parker DM, Tucker-Seeley R. Do minority and poor neighborhoods have higher access to fast-food restaurants in the United States? *Health Place.* 2014;29:10-17.
110. Gordon-Larsen P, Nelson MC, Page P, Popkin BM. Inequality in the built environment underlies key health disparities in physical activity and obesity. *Pediatrics.* 2006;117(2):417-424.
111. Powell LM, Slater S, Chaloupka FJ, Harper D. Availability of physical activity-related facilities and neighborhood demographic and socioeconomic characteristics: A national study. *Am J Public Health.* 2006;96(9):1676-1680.
112. Thomas WI. *On Social Organization and Social Personality: Selected Papers.* Chicago and London: University of Chicago Press; 1966 [1927].
113. Putnam RD. Bowling Alone: America’s Declining Social Capital. In: Crothers L, Lockhart C, eds. *Culture and Politics.* New York, NY: Palgrave Macmillan; 2000.
114. Coleman JS. Social capital in the creation of human capital. *Am J Sociol.* 1988;94:S95-S120.
115. Ralston K. How government policies and regulations can affect dietary choices. In: Frazao E, ed. *America's Eating Habits: Changes and Consequences. Agriculture Information Bull. 750.* Washington, DC: USDA/ERS/Food Rural Econ. Div; 1999.
116. Swinburn BA. Obesity prevention: The role of policies, laws and regulations. *Aust New Zealand Health Policy.* 2008;5(1):12.
117. Drewnowski A, Darmon N. Food choices and diet costs: An economic analysis. *J Nutr.* 2005;135(4):900-904.
118. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. *Lancet.* 2010;376(9748):1261-1271.
119. Hoek J, Gendall P. Advertising and obesity: A behavioral perspective. *J Health Commun.* 2006;11(4):409-423.
120. Merzel C, D'Afflitti J. Reconsidering community-based health promotion: Promise, performance, and potential. *Am J Public Health.* 2003;93(4):557-574.
121. Elder JP, Schmid TL, Dower P, Hedlund S. Community heart health programs: Components, rationale, and strategies for effective interventions. *J Public Health Policy.* 1993;14(4):463-479.

122. Blackburn H. Research and demonstration projects in community cardiovascular disease prevention. *J Public Health Policy.* 1983;4(4):398-421.
123. Ansell LH. Community-Based Health Interventions: Past, Present, and Future. In: Whitman S, Shah AM, Benjamins M, eds. *Urban Health: Combating Disparities with Local Data.* New York: Oxford University Press; 2011.
124. Lalonde M. *A new perspective on the health of Canadians.* Ottawa, ON: Minister of Supply and Services Canada;1974.
125. Dawber TR. *The Framingham Study: The Epidemiology of Atherosclerotic Disease.* Cambridge, Mass.: Harvard University Press; 1980.
126. Rose GA. *The Strategy of Preventive Medicine.* Oxford [England]; New York: Oxford University Press; 1992.
127. McAlister A, Puska P, Salonen JT, Tuomilehto J, Koskela K. Theory and action for health promotion illustrations from the North Karelia Project. *Am J Public Health.* 1982;72(1):43-50.
128. Winkleby MA, Feldman HA, Murray DM. Joint analysis of three U.S. community intervention trials for reduction of cardiovascular disease risk. *J Clin Epidemiol.* 1997;50(6):645-658.
129. Minkler M. Personal responsibility for health? A review of the arguments and the evidence at century's end. *Health Educ Behav.* 1999;26(1):121-140.
130. Frohlich KL, Potvin L. Transcending the known in public health practice: The inequality paradox: The population approach and vulnerable populations. *Am J Public Health.* 2008;98(2):216-221.
131. Economos CD, Irish-Hauser S. Community interventions: A brief overview and their application to the obesity epidemic. *J Law Med Ethics.* 2007;35(1):131-137.
132. Li W, Kelsey JL, Zhang Z, et al. Small-area estimation and prioritizing communities for obesity control in Massachusetts. *Am J Public Health.* 2009;99(3):511-519.
133. Yancey AK, Kumanyika SK, Ponce NA, et al. Population-based interventions engaging communities of color in healthy eating and active living: A review. *Prev Chronic Dis.* 2004;1(1):A09.
134. Black JL, Macinko J. Neighborhoods and obesity. *Nutr Rev.* 2008;66(1):2-20.
135. Letarte L, Pomerleau S, Tchernof A, Biertho L, Waygood EOD, Lebel A. Neighbourhood effects on obesity: Scoping review of time-varying outcomes and exposures in longitudinal designs. *BMJ Open.* 2020;10(3):e034690.
136. Mohammed SH, Habtewold TD, Birhanu MM, et al. Neighbourhood socioeconomic status and overweight/obesity: A systematic review and meta-analysis of epidemiological studies. *BMJ Open.* 2019;9(11):e028238.
137. Kimbro RT, Denney JT. Neighborhood context and racial/ethnic differences in young children's obesity: Structural barriers to interventions. *Soc Sci Med.* 2013;95(Supplement C):97-105.
138. Anderson SE, Whitaker RC. Prevalence of obesity among US preschool children in different racial and ethnic groups. *Arch Pediatr Adolesc Med.* 2009;163(4):344-348.
139. De Milto L. Healthy Kids, Healthy Communities: Supporting community action to prevent childhood obesity. https://www.rwjf.org/en/library/research/2012/02/healthy-kids--healthy-communities.html. Published 2015. Accessed October 30, 2020.

140. UCLA Center for Health Policy Research. California Health Interview Survey. https://healthpolicy.ucla.edu/chis/Pages/default.aspx. Accessed October 31, 2020.
141. Kim Y, Cubbin C, Oh S. A systematic review of neighbourhood economic context on child obesity and obesity-related behaviours. *Obes Rev.* 2019;20(3):420-431.
142. Mitchell TM. *Machine Learning.* New York, NY: McGraw-Hill, Inc.; 1997.
143. Swinburn B. Obesity prevention in children and adolescents. *Child Adolesc Psychiatr Clin N Am.* 2009;18(1):209-223.
144. USDA Food and Nutrition Service. Women, Infants, and Children (WIC). https://www.fns.usda.gov/wic. Accessed October 19, 2020.
145. USDA Food and Nutrition Service. Who Gets WIC and How to Apply. https://www.fns.usda.gov/wic/who-gets-wic-and-how-apply. Updated October 10, 2013. Accessed November 10, 2020.
146. PHFE WIC Program. Data Mining Project. http://apps.phfewic.org/Projects/DataMining.aspx. Accessed October 19, 2020.
147. LAC WIC Data. Los Angeles County WIC Data. http://lawicdata.org/about/. Accessed October 19, 2020.
148. First 5 LA. About Us. https://www.first5la.org/about-us/. Accessed November 1, 2020.
149. Crespi CM, Alfonso VH, Whaley SE, Wang MC. Validity of child anthropometric measurements in the Special Supplemental Nutrition Program for Women, Infants, and Children. *Pediatr Res.* 2012;71(3):286-292.
150. Wang MC, Crespi CM, Jiang LH, et al. Developing an index of dose of exposure to early childhood obesity community interventions. *Prev Med.* 2018;111:135-141.
151. Samuels SE, Craypo L, Boyle M, Crawford PB, Yancey A, Flores G. The California Endowment's Healthy Eating, Active Communities Program: A midpoint review. *Am J Public Health.* 2010;100(11):2114-2123.
152. Kaiser Permanente. HEAL & HEAL Zones. https://community.kp.org/about/program/heal-zones. Accessed September 21, 2020.
153. LA County Department of Public Health. Communities Putting Prevention to Work Initiative. http://www.publichealth.lacounty.gov/arra.htm. Accessed October 19, 2020.
154. Robles B, Jiang L, Prelip M, et al. Notes from the field: Creating a typology of childhood obesity intervention strategies. *Eval Health Prof.* 2020;43(2):131-134.
155. US Census Bureau. Decennial Census of Population and Housing. https://www.census.gov/programs-surveys/decennial-census/about/census-acs.html. Accessed November 1, 2020.
156. University of Missouri Center for Health Policy. Missouri Census Data Center. http://mcdc.missouri.edu/. Accessed November 9, 2020.
157. US Census Bureau. ZIP Code Tabulation Areas (ZCTAs). https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html. Updated August 26, 2020. Accessed November 1, 2020.
158. US Census Bureau. Geographic Areas Reference Manual. https://www.census.gov/programs-surveys/geography/guidance/geographic-areas-reference-manual.html. Updated May 16, 2018. Accessed November 1, 2020.
159. US Census Bureau. *Census 2000 Summary File 1 Technical Documentation.* 2001.

160. Grubesic TH, Matisziw TC. On the use of ZIP Codes and ZIP Code Tabulation Areas (ZCTAs) for the spatial analysis of epidemiological data. *Int J Health Geogr.* 2006;5:58-58.
161. County of Los Angeles Open Data. LA County Zip Codes. https://data.lacounty.gov/Geospatial/LA-County-Zip-Codes/8gwe-vka7. Updated August 29, 2016. Accessed November 1, 2020.
162. US Census Bureau. TIGER/Line® Shapefiles. https://www.census.gov/cgi-bin/geo/shapefiles/index.php. Accessed November 1, 2020.
163. World Health Organization. *Obesity: Preventing and managing the global epidemic.* Geneva: World Health Organization; 2000.
164. Shrewsbury V, Wardle J. Socioeconomic status and adiposity in childhood: A systematic review of cross-sectional studies 1990-2005. *Obesity (Silver Spring).* 2008;16(2):275-284.
165. Carroll-Scott A, Gilstad-Hayden K, Rosenthal L, et al. Disentangling neighborhood contextual associations with child body mass index, diet, and physical activity: The role of built, socioeconomic, and social environments. *Soc Sci Med.* 2013;95(Supplement C):106-114.
166. Caspi CE, Sorensen G, Subramanian SV, Kawachi I. The local food environment and diet: A systematic review. *Health Place.* 2012;18(5):1172-1187.
167. Sawyer ADM, Jones R, Ucci M, Smith L, Kearns A, Fisher A. Cross-sectional interactions between quality of the physical and social environment and self-reported physical activity in adults living in income-deprived communities. *PLOS One.* 2017;12(12):e0188962.
168. Reuben A, Sugden K, Arseneault L, et al. Association of neighborhood disadvantage in childhood with DNA methylation in young adulthood. *JAMA Network Open.* 2020;3(6):e206095-e206095.
169. Nicolaidis S. Environment and obesity. *Metabolism.* 2019;100:153942.
170. Janssen I, Boyce WF, Simpson K, Pickett W. Influence of individual- and area-level measures of socioeconomic status on obesity, unhealthy eating, and physical inactivity in Canadian adolescents. *Am J Clin Nutr.* 2006;83(1):139-145.
171. Mobley LR, Root ED, Finkelstein EA, Khavjou O, Farris RP, Will JC. Environment, obesity, and cardiovascular disease risk in low-income women. *Am J Prev Med.* 2006;30(4):327-332.
172. van Lenthe FJ, Mackenbach JP. Neighbourhood deprivation and overweight: the GLOBE study. *Int J Obes Relat Metab Disord.* 2002;26(2):234-240.
173. Egger G, Swinburn B. An "ecological" approach to the obesity pandemic. *BMJ.* 1997;315(7106):477-480.
174. Los Angeles Times. Mapping L.A.: Neighborhoods. http://maps.latimes.com/neighborhoods/. Accessed October 18, 2020.
175. First 5 LA. Best Start Communities. http://www.first5la.org/index.php?r=site/tag&id=576. Accessed March 23, 2020.
176. Cheadle A, Atiedu A, Rauzon S, et al. A community-level initiative to prevent obesity: Results from Kaiser Permanente's Healthy Eating Active Living Zones Initiative in California. *Am J Prev Med.* 2018;54(5s2):S150-s159.

177. Healthy Eating Active Living Cities Campaign. California. http://www.healnation.com/California. Accessed October 19, 2020.
178. Nobari TZ, Whaley SE, Crespi CM, Prelip ML, Wang MC. Widening socio-economic disparities in early childhood obesity in Los Angeles County after the Great Recession. *Public Health Nutr.* 2018:1-10.
179. Chaparro MP, Anderson CE, Crespi CM, Whaley SE, Wang MC. The effect of the 2009 WIC food package change on childhood obesity varies by gender and initial weight status in Los Angeles County. *Pediatr Obes.* 2019;14(9):e12526.
180. Lopez RP. Neighborhood risk factors for obesity. *Obesity (Silver Spring).* 2007;15(8):2111-2119.
181. Rundle A, Field S, Park Y, Freeman L, Weiss CC, Neckerman K. Personal and neighborhood socioeconomic status and indices of neighborhood walk-ability predict body mass index in New York City. *Soc Sci Med.* 2008;67(12):1951-1958.
182. Eibner C, Sturm R. US-based indices of area-level deprivation: Results from HealthCare for Communities. *Soc Sci Med.* 2006;62(2):348-359.
183. Thomas AJ, Eberly LE, Davey Smith G, Neaton JD. ZIP-code-based versus tract-based income measures as long-term risk-adjusted mortality predictors. *Am J Epidemiol.* 2006;164(6):586-590.
184. U.S. Census Bureau. When to Use 1-year, 3-year, or 5-year Estimates. https://www.census.gov/programs-surveys/acs/guidance/estimates.html. Updated September 10, 2020. Accessed October 22, 2020.
185. U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau. *ZIP Code Tabulation Areas (ZCTAs™).* 2015.
186. Faraway JJ. *Extending the linear model with R : Generalized linear, mixed effects and nonparametric regression models.* Second edition. ed. Boca Raton: CRC Press, Taylor & Francis Group; 2016.
187. Morin R. *One Recession, Two Americas.* Washington, DC: Pew Research Center;2010.
188. Food and Nutrition Service (USDA). Special Supplemental Nutrition Program for Women, Infants and Children (WIC): Revisions in the WIC Food Packages. In. *Code of Federal Regulations (CFR) 7 CFR 246*2008.
189. Zeileis A, Hothorn T. Diagnostic Checking in Regression Relationships. In: R News, 2(3), 7–10.; 2002.
190. *R: A language and environment for statistical computing* [computer program]. Vienna, Austria: R Foundation for Statistical Computing; 2020.
191. Bartok CJ, Ventura AK. Mechanisms underlying the association between breastfeeding and obesity. *Pediatr Obes.* 2009;4(4):196-204.
192. Krieger N, Zierler S, Hogan JW, et al. Geocoding and measurement of neighborhood socioeconomic position: A U.S. perspective. In: Kawachi I, Berkman LF, eds. *Neighborhoods and health.* Oxford: Oxford University Press; 2003.
193. Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual medical spending attributable to obesity: Payer-and service-specific estimates. *Health Aff.* 2009;28(Supplement 1):w822-w831.
194. Whitaker RC, Wright JA, Pepe MS, Seidel KD, Dietz WH. Predicting obesity in young adulthood from childhood and parental obesity. *N Engl J Med.* 1997;337(13):869-873.

195. Wang Y, Lim H. The global childhood obesity epidemic and the association between socio-economic status and childhood obesity. *Int Rev Psychiatry.* 2012;24(3):176-188.
196. Vazquez CE, Cubbin C. Socioeconomic status and childhood obesity: A review of literature from the past decade to inform intervention research. *Curr Obes Rep.* 2020.
197. Freedman DS, Khan LK, Serdula MK, Dietz WH, Srinivasan SR, Berenson GS. The relation of childhood BMI to adult adiposity: The Bogalusa Heart Study. *Pediatrics.* 2005;115(1):22-27.
198. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017-2018. *NCHS Data Brief.* 2020(360):1-8.
199. Joint WHO/FAO Expert Consultation on Diet NatPoCD. *Diet, nutrition and the prevention of chronic diseases: Report of a joint WHO/FAO expert consultation.* Geneva, 2002.
200. Leal C, Chaix B. The influence of geographic life environments on cardiometabolic risk factors: A systematic review, a methodological assessment and a research agenda. *Obes Rev.* 2011;12(3):217-230.
201. Williams DR, Collins C. Racial residential segregation: A fundamental cause of racial disparities in health. *Public Health Rep.* 2001;116(5):404-416.
202. Gustafson A, Hankins S, Jilcott S. Measures of the consumer food store environment: A systematic review of the evidence 2000-2011. *J Community Health.* 2012;37(4):897-911.
203. Alvarado SE. Neighborhood disadvantage and obesity across childhood and adolescence: Evidence from the NLSY children and young adults cohort (1986-2010). *Soc Sci Res.* 2016;57:80-98.
204. Greves Grow HM, Cook AJ, Arterburn DE, Saelens BE, Drewnowski A, Lozano P. Child obesity associated with social disadvantage of children's neighborhoods. *Soc Sci Med.* 2010;71(3):584-591.
205. Burdette AM, Needham BL. Neighborhood environment and body mass index trajectories from adolescence to adulthood. *J Adolesc Health.* 2012;50(1):30-37.
206. Coogan PF, Krishnan S, Cozier YC, et al. Neighborhood socioeconomic status in relation to 10-year weight gain in the Black Women's Health Study. *Obesity (Silver Spring).* 2010;18(10):2064-2065.
207. Do DP, Zheng C. A marginal structural modeling strategy investigating short and long-term exposure to neighborhood poverty on BMI among U.S. black and white adults. *Health Place.* 2017;46:201-209.
208. Elfassy T, Yi SS, Llabre MM, et al. Neighbourhood socioeconomic status and cross-sectional associations with obesity and urinary biomarkers of diet among New York City adults: The Heart Follow-Up Study. *BMJ Open.* 2017;7(12):e018566.
209. Ford PB, Dzewaltowski DA. Neighborhood deprivation, supermarket availability, and BMI in low-income women: A multilevel analysis. *J Community Health.* 2011;36(5):785-796.
210. Rummo PE, Meyer KA, Green Howard A, Shikany JM, Guilkey DK, Gordon-Larsen P. Fast food price, diet behavior, and cardiometabolic health: Differential associations by neighborhood SES and neighborhood fast food restaurant availability in the CARDIA study. *Health Place.* 2015;35:128-135.

211. Kimbro RT, Sharp G, Denney JT. Home and away: Area socioeconomic disadvantage and obesity risk. *Health Place.* 2017;44:94-102.
212. Kwarteng JL, Schulz AJ, Mentz GB, Israel BA, Perkins DW. Independent effects of neighborhood poverty and psychosocial stress on obesity over time. *J Urban Health.* 2017;94(6):791-802.
213. Kwarteng JL, Schulz AJ, Mentz GB, Israel BA, Shanks TR, Perkins DW. Neighbourhood poverty, perceived discrimination and central adiposity in the USA: Independent associations in a repeated measures analysis. *J Biosoc Sci.* 2016;48(6):709-722.
214. Lovasi GS, Jacobson JS, Quinn JW, Neckerman KM, Ashby-Thompson MN, Rundle A. Is the environment near home and school associated with physical activity and adiposity of urban preschool children? *J Urban Health.* 2011;88(6):1143-1157.
215. Lippert AM, Evans CR, Razak F, Subramanian SV. Associations of continuity and change in early neighborhood poverty with adult cardiometabolic biomarkers in the United States: Results from the National Longitudinal Study of Adolescent to Adult Health, 1995-2008. *Am J Epidemiol.* 2017;185(9):765-776.
216. Mendez DD, Gary-Webb TL, Goode R, et al. Neighborhood factors and six-month weight change among overweight individuals in a weight loss intervention. *Prev Med Rep.* 2016;4:569-573.
217. Mujahid MS, Diez Roux AV, Borrell LN, Nieto FJ. Cross-sectional and longitudinal associations of BMI with socioeconomic characteristics. *Obes Res.* 2005;13(8):1412-1421.
218. Mujahid MS, Diez Roux AV, Shen M, et al. Relation between neighborhood environments and obesity in the Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol.* 2008;167(11):1349-1357.
219. Powell-Wiley TM, Cooper-McCann R, Ayers C, et al. Change in neighborhood socioeconomic status and weight gain: Dallas Heart Study. *Am J Prev Med.* 2015;49(1):72-79.
220. Powell-Wiley TM, Ayers C, Agyemang P, et al. Neighborhood-level socioeconomic deprivation predicts weight gain in a multi-ethnic population: Longitudinal data from the Dallas Heart Study. *Prev Med.* 2014;66:22-27.
221. Ruel E, Reither EN, Robert SA, Lantz PM. Neighborhood effects on BMI trends: Examining BMI trajectories for Black and White women. *Health Place.* 2010;16(2):191-198.
222. Wang MC, Kim S, Gonzalez AA, MacLeod KE, Winkleby MA. Socioeconomic and food-related physical characteristics of the neighbourhood environment are associated with body mass index. *J Epidemiol Community Health.* 2007;61(6):491-498.
223. Messer LC, Laraia BA, Kaufman JS, et al. The development of a standardized neighborhood deprivation index. *J Urban Health.* 2006;83(6):1041-1062.
224. Carroll-Scott A, Gilstad-Hayden K, Rosenthal L, et al. Associations of neighborhood and school socioeconomic and social contexts with body mass index among urban preadolescent students. *Am J Public Health.* 2015;105(12):2496-2502.
225. Lovasi GS, Schwartz-Soicher O, Quinn JW, et al. Neighborhood safety and green space as predictors of obesity among preschool children from low-income families in New York City. *Prev Med.* 2013;57(3):189-193.

226. Nelson MC, Gordon-Larsen P, Song Y, Popkin BM. Built and social environments associations with adolescent overweight and activity. *Am J Prev Med.* 2006;31(2):109-117.
227. Shih M, Dumke KA, Goran MI, Simon PA. The association between community-level economic hardship and childhood obesity prevalence in Los Angeles. *Pediatr Obes.* 2013;8(6):411-417.
228. Kowaleski-Jones L, Wen M. Community and child energy balance: Differential associations between neighborhood environment and overweight risk by gender. *Int J Environ Health Res.* 2013;23(5):434-445.
229. McTigue KM, Cohen ED, Moore CG, Hipwell AE, Loeber R, Kuller LH. Urban neighborhood features and longitudinal weight development in girls. *Am J Prev Med.* 2015;49(6):902-911.
230. Ohri-Vachaspati P, Lloyd K, Delia D, Tulloch D, Yedidia MJ. A closer examination of the relationship between children's weight status and the food and physical activity environment. *Prev Med.* 2013;57(3):162-167.
231. Rossen LM. Neighbourhood economic deprivation explains racial/ethnic disparities in overweight and obesity among children and adolescents in the U.S.A. *J Epidemiol Community Health.* 2014;68(2):123-129.
232. Saelens BE, Sallis JF, Frank LD, et al. Obesogenic neighborhood environments, child and parent obesity: The Neighborhood Impact on Kids Study. *Am J Prev Med.* 2012;42(5):e57-e64.
233. Voorhees CC, Catellier DJ, Ashwood JS, et al. Neighborhood socioeconomic status and non school physical activity and body mass index in adolescent girls. *J Phys Act Health.* 2009;6(6):731-740.
234. Chang VW. Racial residential segregation and weight status among US adults. *Soc Sci Med.* 2006;63(5):1289-1303.
235. Inagami S, Cohen DA, Finch BK, Asch SM. You are where you shop: Grocery store locations, weight, and neighborhoods. *Am J Prev Med.* 2006;31(1):10-17.
236. Boardman JD, Saint Onge JM, Rogers RG, Denney JT. Race differentials in obesity: The impact of place. *J Health Soc Behav.* 2005;46(3):229-243.
237. Robert SA, Reither EN. A multilevel analysis of race, community disadvantage, and body mass index among adults in the US. *Soc Sci Med.* 2004;59(12):2421-2434.
238. National Research Council. *Health Performance Measurement in the Public Sector: Principles and Policies for Implementing an Information Network.* Washington, DC: The National Academies Press; 1999.
239. dos Santos BS, Steiner MTA, Fenerich AT, Lima RHP. Data mining and machine learning techniques applied to public health problems: A bibliometric analysis from 2009 to 2018. *Computers & Industrial Engineering.* 2019;138:106120.
240. Ahmadi H, Gholamzadeh M, Shahmoradi L, Nilashi M, Rashvand P. Diseases diagnosis using fuzzy logic methods: A systematic and meta-analysis review. *Computer Methods and Programs in Biomedicine.* 2018;161:145-172.
241. Benedetto U, Dimagli A, Sinha S, et al. Machine learning improves mortality risk prediction after cardiac surgery: Systematic review and meta-analysis. *J Thorac Cardiovasc Surg.* 2020.

242. Bellinger C, Mohomed Jabbar MS, Zaïane O, Osornio-Vargas A. A systematic review of data mining and machine learning for air pollution epidemiology. *BMC Public Health.* 2017;17(1):907.
243. Dugan TM, Mukhopadhyay S, Carroll A, Downs S. Machine Learning Techniques for Prediction of Early Childhood Obesity. *Applied Clinical Informatics.* 2015;6(3):506-520.
244. Hammond R, Athanasiadou R, Curado S, et al. Predicting childhood obesity using electronic health records and publicly available data. *PLOS One.* 2019;14(4):e0215571.
245. Montañez CAC, Fergus P, Hussain A, et al. Machine learning approaches for the prediction of obesity using publicly available genetic profiles. Paper presented at: 2017 International Joint Conference on Neural Networks (IJCNN); 14-19 May 2017, 2017.
246. Zhang S, Tjortjis C, Zeng X, Qiao H, Buchan I, Keane J. Comparing data mining methods with logistic regression in childhood obesity prediction. *Inf Syst Front.* 2009;11(4):449-460.
247. Luo W, Nguyen T, Nichols M, et al. Is demography destiny? Application of machine learning techniques to accurately predict population health outcomes from a minimal demographic dataset. *PLOS ONE.* 2015;10(5):e0125602.
248. Rose S. Mortality risk score prediction in an elderly population using machine learning. *Am J Epidemiol.* 2013;177(5):443-452.
249. Murray CJL. Towards good practice for health statistics: Lessons from the Millennium Development Goal health indicators. *Lancet.* 2007;369(9564):862-873.
250. U.S. Census Bureau. About the American Community Survey. https://www.census.gov/programs-surveys/acs/about.html. Updated October 16, 2020. Accessed May 3, 2018.
251. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat.* 1996;58(1):267-288.
252. Doupe P, Faghmous J, Basu S. Machine learning for health services researchers. *Value Health.* 2019;22(7):808-815.
253. Botchkarev A. Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. *arXiv preprint arXiv:180903006.* 2018.
254. Kuhn M. Building predictive models in R using the caret package. *Journal of Statistical Software.* 2008;28(5):1-26.
255. Seliske L, Norwood TA, McLaughlin JR, Wang S, Palleschi C, Holowaty E. Estimating micro area behavioural risk factor prevalence from large population-based surveys: a full Bayesian approach. *BMC Public Health.* 2016;16:478.
256. Cho WKT, Manski CF. Cross level/ecological inference. In: Box-Steffensmeier JM, Brady HE, Collier D, eds. *Oxford handbook of political methodology.* Oxford University Press; 2008:547-569.
257. Yan J, Liu L, Zhu Y, Huang G, Wang PP. The association between breastfeeding and childhood obesity: A meta-analysis. *BMC Public Health.* 2014;14(1):1267.
258. Robinson PL, Dominguez F, Teklehaimanot S, Lee M, Brown A, Goodchild M. Does distance decay modelling of supermarket accessibility predict fruit and vegetable intake by individuals in a large metropolitan area? *J Health Care Poor Underserved.* 2013;24(1 Suppl):172-185.

259. Institute of Medicine. *Accelerating Progress in Obesity Prevention: Solving the Weight of the Nation.* Washington, DC: The National Academies Press;2012.
260. Waters E, de Silva-Sanigorski A, Hall BJ, et al. Interventions for preventing obesity in children. *Cochrane Database Syst Rev.* 2011(12):1-212.
261. Bell LM, Byrne S, Thompson A, et al. Increasing Body Mass Index z-Score Is Continuously Associated with Complications of Overweight in Children, Even in the Healthy Weight Range. *The Journal of Clinical Endocrinology & Metabolism.* 2007;92(2):517-522.
262. Centers for Disease Control and Prevention. NCCDPHP: Community Health. https://www.cdc.gov/nccdphp/dch/programs/past_programs/index.htm. Published 2017. Accessed November 4, 2020.
263. Arteaga SS, Loria CM, Crawford PB, et al. The Healthy Communities Study: Its rationale, aims, and approach. *Am J Prev Med.* 2015;49(4):615-623.
264. James G, Witten D, Hastie T, Tibshirani R. *An introduction to statistical learning.* Vol 112. New York, NY: Springer; 2013.
265. Zou H, Hastie T. Regularization and variable selection via the Elastic Net. *J R Stat Soc Series B Stat.* 2005;67(2):301-320.
266. Strauss WJ, Nagaraja J, Landgraf AJ, et al. The longitudinal relationship between community programmes and policies to prevent childhood obesity and BMI in children: The Healthy Communities Study. *Pediatr Obes.* 2018;13(S1):82-92.
267. Solar O, Irwin A. *A conceptual framework for action on the social determinants of health.* Geneva, Switzerland: WHO;2010.
268. Bleich SN, Segal J, Wu Y, Wilson R, Wang Y. Systematic review of community-based childhood obesity prevention studies. *Pediatrics.* 2013;132(1):e201-210.
269. Hupkens CLH, Knibbe RA, Drop MJ. Social class differences in food consumption: The explanatory value of permissiveness and health and cost considerations. *Eur J Public Health.* 2000;10(2):108-113.
270. Vogel C, Lewis D, Ntani G, et al. The relationship between dietary quality and the local food environment differs according to level of educational attainment: A cross-sectional study. *PLOS One.* 2017;12(8):e0183700.
271. Adams J, Mytton O, White M, Monsivais P. Why are some population interventions for diet and obesity more equitable and effective than others? The role of individual agency. *PLOS Med.* 2016;13(4):e1001990.
272. Monasta L, Batty GD, Macaluso A, et al. Interventions for the prevention of overweight and obesity in preschool children: A systematic review of randomized controlled trials. *Obes Rev.* 2011;12(5):e107-e118.
273. Giugliani ERJ, Horta BL, Loret de Mola C, Lisboa BO, Victora CG. Effect of breastfeeding promotion interventions on child growth: A systematic review and meta-analysis. *Acta Paediatrica.* 2015;104(S467):20-29.
274. Koplin JJ, Kerr JA, Lodge C, et al. Infant and young child feeding interventions targeting overweight and obesity: A narrative review. *Obes Rev.* 2019;20(S1):31-44.
275. Brug J, van Lenthe FJ, Kremers SP. Revisiting Kurt Lewin: How to gain insight into environmental correlates of obesogenic behaviors. *Am J Prev Med.* 2006;31(6):525-529.

276. Institute of Medicine. *Evaluating Obesity Prevention Efforts: A Plan for Measuring Progress.* Washington, DC: The National Academies Press;2013.