The Built Environment, Obesity and Walking

Ruizhu Huang

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Anne V. Moudon, Chair

Brian E. Saelens

Qing Shen

Program Authorized to Offer Degree:

Urban Design and Planning

University of Washington

Abstract

The Built Environment, Obesity and Walking

Ruizhu Huang

Chair of the Supervisory Committee:
Professor Anne Vernez Moudon
Urban Design & Planning

This dissertation focuses on the influence of the built environment on the geographic distribution of walking behavior and the obese epidemic. Fundamentally, obesity results from energy imbalance. The study aims to help planners understand how to change neighborhood environment in order to reconstruct energy balance: decreasing energy intake by supplying healthy food environment and increasing energy output by introducing physical activity supporting environment.

Neighborhood-level factors are hypothesized to influence obesity and walking behavior via exposure and access mechanisms (Feng et al., 2010; Moudon et al., 2007; Saelens and Handy, 2008). The hypothesis raises the following questions for this dissertation to answer: Does obesity cluster in space? Where and why does obesity cluster? How much walking occurs around home? What home neighborhood built environment characteristics are indeed associated with walking? How may individuals change their walking behavior after introduction of light rail?

By answering the questions above, the dissertation presents new evidence for future prevention studies: The geographic distribution of obesity is not random and related to neighborhood property value and residential density. Taking Seattle King County as an

example, this area presents a north to south obesity gradient, with the northern part being less obese, richer and more densely populated and the southern part being more obese, lower income, and less densely populated. In terms of walking, the evidence that over half of walking occurs within the home neighborhood emphasizes the importance of the role of the home neighborhood in promoting physical activity. Higher residential and job density, which correlate with high street intersection, sidewalk and fitness density, would effectively support walking behavior in home neighborhoods. Also, after the introduction of light rail, people tend to walk more in station areas, which implies that rail transit may help increase the number of potential customers for retail and service near station areas.

This dissertation has limitations that point to areas of future research. The circular shapes used to detect obesity clusters may not really represent the boundary of a neighborhood. In reality, neighborhood boundaries have complex shapes (Kulldorff et al., 2006; Patil and Taillie, 2004; K Takahashi et al., 2008). More research needs to be done on incorporating detection methods with neighborhood boundaries usually separated by natural obstacles such as lakes or rivers. Finally, more natural experiments with longer assessment period need to be carried out to compare walking behavior in free-living daily activities before and after interventions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The author wishes to express sincere appreciation to the Urban Form Lab, the Department of Urban Design and Planning and the Interdisciplinary PhD program of Urban Design and Planning at the University of Washington for their long-term support. My deep thanks go to my dissertation supervisor Professor Anne V. Moudon for her support and guidance during my five-year PhD education. Without her help, I would have never been able to obtain a PhD degree from the University of Washington. I believe what I learned from Prof. Anne V. Moudon will be cherished throughout my whole life. To the rest of my dissertation committee – Brian Saelens, Qing Shen, Adrian Dobra and Yinhai Wang – I am grateful for their generous sharing of ideas, time, and expertise.

I would like to thank Professor Adam Drewnowski and Professor Brian Saelens for allowing me to use data from the Seattle Obesity Study (SOS) and the Effect of Light Rail Transit on Physical Activity: A Natural Experiment respectively for my dissertation research. I want to thank Dr. Andrea Cook and Dr. Chuan Zhou for their support in data analysis.

I also want to thank my colleagues and friends in the Urban Form Lab, University of Washington, for being supportive during my PhD study. Thanks goes to Dr. Phil Hurvitz for assisting me with various questions, to Dr. Bumjoon Kang for sharing his research experience, to Mr. Orion Stewart for helping on my writing, to Mr. Jason Scully and Mr. Eric Howard for their support in my research.

Data and parts of the data development for this research came from the Seattle Obesity Study (SOS) funded by the National Institute of Health (NIH, R01 DK076608) through the Center for Obesity Research (UW-COR), the Effect of Light Rail Transit on Physical Activity: A Natural Experiment (NIH/NHLBI R01HL091881 and by the Washington Transportation Center TransNow Research Project Agreement No. 61-7318) and the Urban Form Lab (UFL) at the University of Washington.

DEDICATION

To my parents, my sister and my wife

Chapter 1 INTRODUCTION

Promoting regular involvement in walking is one strategy for overcoming the public health burden of physical inactivity and obesity in developed countries (Eyler et al., 2003). Thus, understanding the neighborhood determinants of walking and obesity has emerged as an important topic in epidemiologic research (Brownson et al., 2005) on neighborhoods and health (Kawachi and Berkman, 2003). Most scientists believe that the obesity epidemic is being driven by changes in eating and exercise behaviors caused by changes in the built environment. Nevertheless, the built environments around us are complex physical, economic, and sociocultural systems that function to influence and constrain people's behaviors through an array of processes. In reality our behaviors are not determined by environmental conditions, but environments work by limiting or enabling choices and by rewarding or inhibiting specific behaviors. Hence, neighborhood built environments can indirectly influence health by affecting knowledge, attitudes, beliefs, and behaviors.

Neighborhood built environment characteristics are hypothesized to influence obesity via exposure and access mechanisms. For example, the exposure to fast food restaurants and convenience stores is associated with higher weight status (Spence, Cutumisu et al. 2009). Neighborhood access to supermarkets and groceries is associated with lower body weight (Morland, Diez Roux et al. 2006; Rundle, Richards et al. 2009). Furthermore, the lack of stores and restaurants in residential neighborhood may force people to travel by car, leading to low levels of daily physical activity; while neighborhoods with activity supportive characteristics such as high residential density and mixed land uses have been found to be associated with more physical activity and lower body weight (Papas, Alberg et al. 2007; Feng, Curriero et al. 2010).

This dissertation includes three essays and aims to disentangle the hidden effects of the built environment on obesity and walking behaviors. The first essay (chapter 2) used spatial cluster detection methods to identify obesogenic neighborhoods and the associated key features of the built environment. The second essay (chapter 3) explored the relationship between home built environment and people's walking bouts frequency around home. The third essay (chapter 4) investigated longitudinal change in walking around light rail stations.

The subject-based geographic buffer neighborhoods used in chapter 2 and chapter 3 are considered to be behaviorally more relevant than administrative boundaries to describe neighborhoods, because they describe the contextual environment for each individual. Administrative boundaries have frequently been used because data sources are conveniently available at the scales of census tract and county. They have been criticized for not being experientially relevant (Moudon, Lee et al. 2006).

The longitudinal study presented in chapter 4 fills a gap as the majority of current studies have use cross-sectional designs. Cross-sectional studies examine the relationship between conditions (e.g., physical activity behaviors) and other variables of interest in a defined population at a single point in time. The reliance on cross-sectional designs can be explained by the nature of the phenomenon of interest. The built environment is relatively fixed in any particular area, changing little or slowly over time. If the built environment in a place changes more radically, residents are often displaced (Handy 2004). The use of simple conceptual models that assume that a walkable environment has a causal effect on walking levels with cross-sectional research designs leaves many unanswered questions about the causal mechanisms involved. In particular, the possibility of "self-selection" must be addressed. It is possible, for example, that individuals who prefer to walk more choose to live in neighborhoods that are more conducive to walking. If researchers do not properly account for the choice of neighborhood, their empirical results will be biased in the sense that features of the built environment may appear to influence activity, while in reality the subjects' decision to walk are the primary impetus for their walking (Feng et al., 2010; Papas et al., 2007).

Objective measure of walking bouts derived from GPS, accelerometer as outcome in chapter 3 and chapter 4 can provide accurate information on when, where and how much walking occurs. Most current studies seeking to find out where people walk have been limited by their use of self-reported data on the behavior. Self-report data are imprecise. For example, responses to the International Physical Activity Questionnaire (IPAQ) typically over-estimate walking frequency (Lee et al., 2011; Rzewnicki et al., 2003) while transportation surveys focusing on motorized trips tend to underreport walking trips (Stopher and Greaves, 2007) Self-reported data also cannot provide precise information about locations where walking occurs (Kang et al., 2013).

The importance of the home neighborhood was tested by examining the concentration of walking around home in chapter 3. The home neighborhood is critically important because it is the origin and ultimate destination from and to which the activity space is defined. Activity space is defined as "the subset of all urban locations within which the individual has direct contact as the result of day-to-day activities. It is characterized as a surface (both discrete and continuous) descriptive of the intensity of actual spatial behavior over portions of the action space" (Horton and Reynolds 1971).

Multicollinearity issues in built environment characteristics were addressed in chapter 3. Many studies find significant association between individual neighborhood characteristics and walking, but ignore the fact that some neighborhood environment variables are highly correlated. The issue may lead to erroneous estimates of coefficients, large standard errors in the related independent variables and spurious significant results, though it does not reduce the predictive power of the model as a whole.

Chapter 2 : THE SPATIAL CLUSTERING OF OBESITY: DO THE SOCIAL AND BUILT ENVIRONMENTS MATTER?

ABSTRACT

Obesity rates in the US show distinct geographic patterns. This study used spatial cluster detection methods to identify obesogenic neighborhoods and the associated key features of the built environment. The Seattle Obesity Study (SOS) provided data on self-reported height, weight, and socio-demographic characteristics of 1602 King County adults. Home addresses were geocoded. Spatially continuous values of local environment features (SmartMaps) were constructed for seven environmental variables. Spatial clusters of significantly high or low BMI were identified, based on Anselin's Local Moran's I and spatial scan statistics, adjusting for individual-level demographic and SES covariates. The clustering of body mass index (BMI) by geographic location, as observed using the two methods, was attenuated after adjusting for residential density and for residential property values obtained from tax rolls and used here as a measure of neighborhood deprivation or wealth. The spatial concentration of obesity was wholly explained by neighborhood composition and socioeconomic characteristics. These characteristics may serve to more precisely locate obesity prevention and intervention programs.

INTRODUCTION

The socio-demographic profile of obesity remains a subject of debate in the U.S. (Chang and Christakis, 2005; Wang and Beydoun, 2007). Income and gender stand out as explanatory factors, with some studies finding an inverse relation between obesity and socioeconomic status for women, but observing no clear patterns for men (Drewnowski and Specter, 2004; Wang and Beydoun, 2007). Social networks and family appear to further explain the grouping of obesity (Cameron et al., 2011; Christakis and Fowler, 2007; Gallos et al., 2012). As well, geographic disparities in obesity are noted by state (Centers for Disease Control and Prevention, n.d.), county (Mokdad et al., 2003), ZIP Code (Drewnowski et al., 2007), Census tract (Ludwig et al., 2011), community boundary (Shih et al., 2013), and neighborhood (Schuurman et al., 2009).

Spatial cluster detection is a useful tool in disease surveillance to identify areas of elevated risk and to generate hypotheses about disease etiology. If a disease clustering is detected then the areas of high residual risk will lead to an 'excess' of cases in those areas-such a collection of cases is defined as a cluster. With this definition, a cluster may be over a very large geographical area. The typical data structure for assessments of spatial health patterns involves a collection of locations of incident events over a particular period for a given study area. For example, one might document the residential locations for children diagnosed with acute lymphocytic leukemia in a given year. Common questions relating to the clustering of health events include: Do cases tend to occur near other cases (perhaps suggesting an infectious agent)? Does a particular area within the study region seem to contain a significant excess of observed cases (perhaps suggesting an environmental risk factor)? Where are the most unusual collections of cases (the most likely clusters)? (L A Waller and Gotway, 2004)

Most scientists believe that the obesity epidemic is being driven by changes in eating and exercise behaviors, which are, in turn, caused by changes in the environment. Environments are complex physical, economic, and sociocultural systems that function to influence and constrain people's behaviors through an array of processes. An individual's behavior is not determined by environmental conditions, but environments work by limiting or enabling choices and by rewarding or punishing specific behaviors. Neighborhood characteristics are aspects of a person's geographic location that can indirectly influence health by affecting knowledge, attitudes, beliefs, and behaviors. A number of risk factors for poor health that often characterize disadvantaged neighborhoods, including high levels of poverty, poor working conditions, high unemployment, discrimination, and limited social capital (defined as community infrastructure that supports education, health, and welfare)(Mobley, Finkelstein et al. 2004; Monda and Popkin 2005).When applied to populations of individuals, patterns of behavior and health will emerge and covary with environmental characteristics. While patterns of covariation will not establish cause and effect, they may be used to generate testable hypotheses(Schlundt, Hargreaves et al. 2006). Essentially, the purpose of examine the clustering of health outcomes and health behaviors is to learn the extent to which some neighborhoods are healthier than others and what neighborhood characteristics are associated with healthier and less healthy places to live.

Identifying if, and explaining how, obesity clusters spatially at the finest scale, and specifically at the local level, will serve to directly target prevention and intervention strategies (Penney et al., 2013). To this end, studies linking obesity with selected features of the built environment have provided evidence that exposure and access to healthy food and the quality of the neighborhood environment influenced diet and physical activity behaviors (Block et al., 2011; Feng et al., 2010; Gordon-Larsen et al., 2006; Hill Peters, J C, 1998; Papas et al., 2007; Rehm et al., 2012).Hence, one might expect to find that spatially distinct patterns of BMI would be associated with some salient underlying features of the built environment at the local level (Feng et al., 2010; Papas et al., 2007). Further comprehensive spatial analysis including environmental determinants of both diet and physical activity will help to disentangle the complex role that the local environment plays in obesity, and more generally in the health of communities (Penney et al., 2013).

The spatial distribution of obesity has been mainly studied at some level of geographic aggregation whether obesity data were obtained at the individual level or not. Most studies used multilevel models with a spatial correlation structure (Diez-Roux, 1998). For these models, the selection of appropriate spatial units of analysis becomes critically important, because the degree of clustering can differ by geographic scale and can affect results (Michimi and Wimberly, 2010; L A Waller and Gotway, 2004). Furthermore, aggregated data are spatially dependent or autocorrelated, leading to biased estimates.

This study focused on whether individuals exhibiting similar body weights clustered within the same local neighborhood environment. Individual home addresses served to accurately locate each participant. The clustering of residuals from regression models was used as the outcome of the analyses in order to specify error dependence structure for inference. Two cluster detection techniques were employed, Anselin's Local Moran' I (Crépey and Barthélemy, 2007; Greene et al., 2006; Kitron and Kazmierczak, 1997; Schuurman et al., 2009) and spatial scan statistic (Gesink et al., 2011; Huang et al., 2009; Kulldorff et al., 1997; Omer et al., 2008; Sabel et al., 2003), which have previously been applied to a range of spatial epidemiologic studies examining geographic patterns in diseases and health behaviors. The results from the two methods are compared to take into account possible bias due to differences in the assumptions made and the spatial patterns generated (Jacquez and Greiling, 2003).

The hypothesis was that spatial clustering by Body Mass Index (BMI) would be explained by the characteristics of the respondents' home neighborhood and by specific features of the built environment, measured using a SmartMaps technique.

METHODS

Study population

Data came from the Seattle Obesity Study (SOS), a population-based study of social disparities, diet quality, and health. A stratified sampling scheme was used to ensure adequate representation by income range and race/ethnicity. King County zip codes with high percentages of household incomes <$35,000, African-Americans or Hispanics were over-sampled. Following standard procedures, randomly generated telephone numbers were matched with residential addresses using commercial databases. A pre-notification letter was mailed out to alert household members that they had been selected at random to participate in a study conducted by the University of Washington School of Public Health. Telephone calls were placed mostly in the afternoons and evenings with up to 13 follow ups to reach the pre-selected numbers. Once the household was contacted, an adult member of the household was randomly selected to be the survey respondent. Exclusion criteria were age less than 18 years, discordance between telephone numbers and addresses obtained from the vendor compared to those obtained from each respondent in beginning of the telephone survey, and cell phone numbers. A 20 min telephone survey was then administered by trained and computer assisted interviewers to 2,001 survey study participants. Administered between October 2008 and March 2009, the survey was approved by the Institutional Review Board at University of Washington.

SOS participants were representative of the population in King County in terms of race, ethnicity, income and household size; and similar to the Behavioral Risk Factor Surveillance System (BRFSS) 2008 sample in terms of age and gender ("Washington State Behavioral Risk Factor Surveillance System [BRFSS]. Questionnaire, forms A and B," 2008). Both response rates (32%) and cooperation rate (97%) were higher than the BRFSS 2008 for King County (29% and 70% respectively). The survey provided data on the respondents' residential address, demographic characteristics (age, race, and gender), socioeconomic status (education and income), and self-reported height and weight (used to calculate BMI).

The addresses of SOS respondents were geocoded to the centroid of the home parcel using the 2008 King County Assessor parcel data. Geocoding followed standard methods in ArcGIS, version 9.3.1. Address records that failed the automatic geocoding (30%, using a 100% match score) were manually matched using a digital map environment with annotated layers from the reference data, augmented by online resources such as GoogleMaps. Each home point was double-checked by a separate technician for plausibility (i.e., that the parcel was designated for residential land use) and accuracy (i.e., that the location was on the correct parcel). Total 1870 home addresses were successfully geocoded on parcels. Participants with missing age, race, income, and education data were excluded. In all, 399 participants were excluded, resulting in a final study population of 1,602. The geographic distribution of residential location is shown in Figure 2-1.

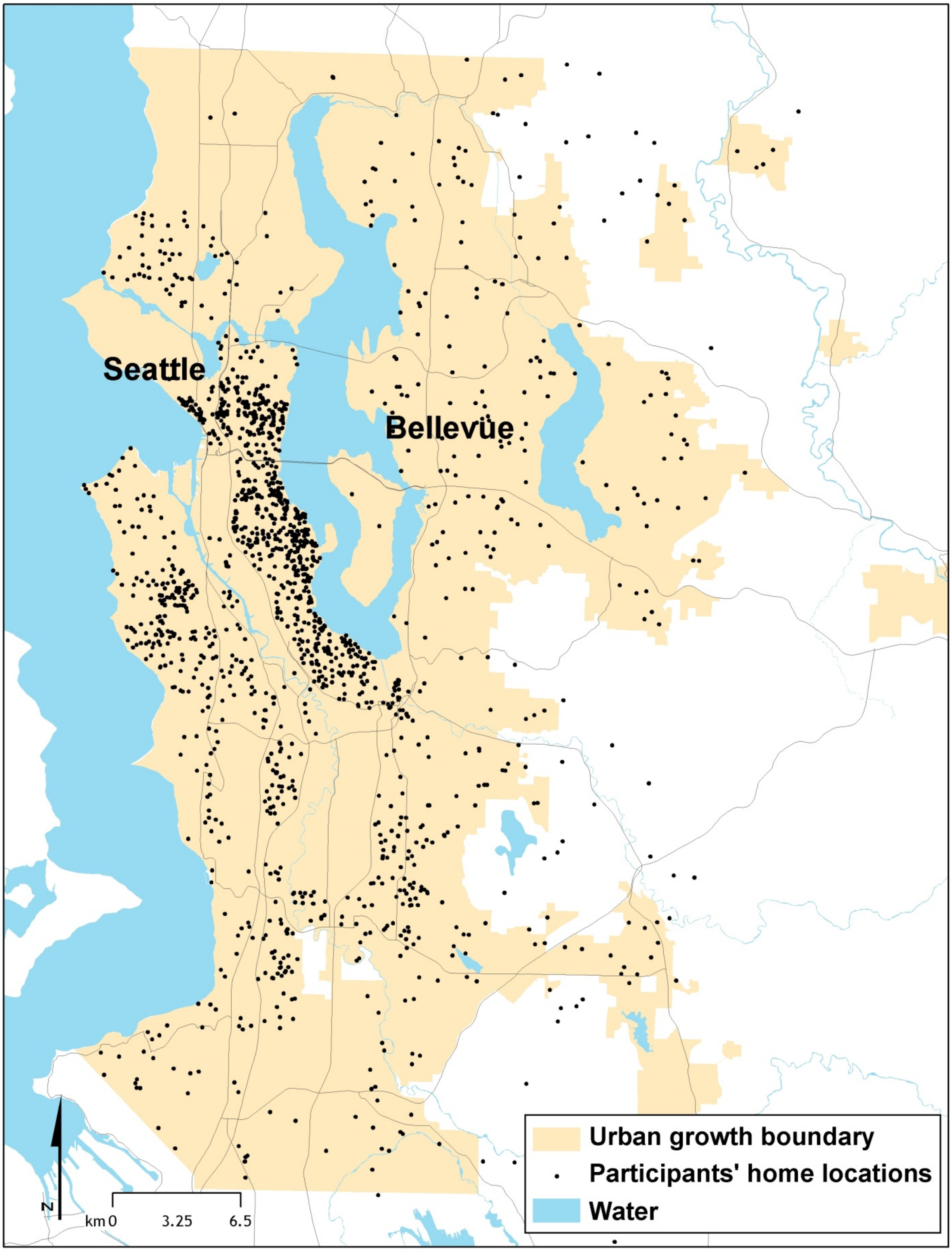

Figure 2-1 The geographic distribution of participants' home

Neighborhood environment

Objective data of the neighborhood built environment came from the King County assessor and GIS Center, from the Department of Public Health of Seattle and King County, and the Urban Form Lab, as described before (Drewnowski et al., 2012; Jiao et al., 2010; Moudon et al., 2011a). All data were captured at the parcel (tax lot) or street segment level. Seven neighborhood environment variables were selected because they had previously been associated with obesity. Residential density measured neighborhood composition (Ewing et al., 2008; Feng et al., 2010; Forsyth et al., 2007; Papas et al., 2007); and residential property values represented neighborhood socioeconomic status (SES) (Moudon et al., 2011a; Rehm et al., 2012). The food environment was captured by the density of supermarkets and grocery stores (Feng et al., 2010; Papas et al., 2007; Rundle et al., 2009) and of fast food and quick service restaurants (Feng et al., 2010; Li et al., 2009; Papas et al., 2007). The active living variable was the percent of park land in the neighborhood (Feng et al., 2010; Papas et al., 2007), and the transportation environment included traffic volumes and intersection density (Frank et al., 2007; Moudon et al., 2007).

The seven variables were processed into SmartMaps using focal raster processing in GIS (Hurvitz and Moudon, 2012; Hurvitz et al., 2014). First, the study area was overlaid with a 30x30 m grid, previously shown to represent urban and suburban parcels with sufficient spatial fidelity (Moudon et al., 2011b). Each grid cell was assigned a neighborhood, defined within an 833 m radius (a 10-minute walking distance) of the cell centroid. The mean value of the environmental variables was calculated for the neighborhood around each grid cell using ArcGIS (version 10.0) Spatial Analyst Extension. Environmental variables were thus converted to a continuous surface across the study area with mean neighborhood values available at each grid point, and then linked spatially to our study participants' geocoded addresses for analyses.

Spatial cluster detection methods

Spatial clusters of participants with high or low BMI values were detected using Anselin's Local Moran's I and a spatial scan statistic. The Local Moran's I (Anselin, 1995; Getis and Ord, 1996) assumes that each participant indexed by *i* has a continuous outcome Y_i, with a spatial location (c_i,d_i). The statistic of spatial cluster at a location (c_i,d_i) is given as:

$$I_i = \frac{Y_i - \overline{Y}}{s_i} \sum_{j=1, j \neq i}^{n} w_{i,j} \frac{Y_j - \overline{Y}}{s_j}$$

where $\overline{Y}$ is the mean of the Y_i's in the population; $w_{i,j}$ is the spatial weight between individual i and j; and s is the square root of the standardized variance of the Y_i's. Spatial weights can be defined as the inverse distance or distance band between individuals. To be comparable to the results of spatial scan statistic, spatial weights $w_{i,j}$ were defined using a fixed distance (2 miles) band method, assigning 1 for neighbors residing within a specified distance and 0 otherwise. Two miles corresponded to the median distance between home and the primary supermarket reported to be used by the SOS population (Drewnowski et al., 2013); the distance was considered small enough to represent a neighborhood and large enough to provide sufficient power to detect clusters.

The statistical significance of the Local Moran's I statistic was tested using conditional permutation (Anselin et al., 2006), an appropriate method for BMI data, which is usually not normally distributed . The conditional permutation approach fixes all observed locations and then randomly permutes all observed outcome values to form a permuted dataset assuming no relationship between location and outcome (i.e. no clustering). For each of these permuted datasets, a Local Moran's I test statistic is calculated to form a reference distribution under the NULL hypothesis of no clustering. Using this permuted Moran's I test statistic distribution, the following "pseudo" p-value can be calculated:

$$pval_i = \frac{M_i + 1}{Nperm + 1}$$

where M_i is the number of instances where a permuted test statistic is equal to or greater than the observed value (for positive Local Moran' I index) or less or equal to the observed value (for negative Local Moran's I index); and Nperm is the number of permuted test statistics. To control for multiple comparisons of testing all potential clusters (Anselin et al., 2006), we used a conservative 0.001 significance level instead of the standard 0.05.

The spatial scan statistic calculates a likelihood ratio test (LRT) statistic evaluating whether observations within a cluster have higher (or lower) outcome values relative to all observations outside of a cluster. A potential cluster is a circular region around each participant's home. The maximum search window was set to 2 miles. The detection of a significant cluster of high/low BMI indicates that respondents inside the cluster have a significantly higher/lower likelihood of having high/low BMI, compared with respondents living outside of the cluster.

The spatial scan statistic uses a similar conditional permutation approach to calculate statistical significance as described for Moran's I. However, it handles the multiple comparison problem more explicitly. For each permuted dataset, LRT statistics are calculated for all potential clusters (across all locations for all varying radius sizes). One then calculates the maximum LRT for that permuted dataset forming a distribution of maximum LRT statistics. Using this permuted maximum LRT statistics dataset one compares the maximum observed LRT to the permuted maximum LRT's and calculates a pseudo p-value,

$$pval = \frac{\sum_{p=1}^{Nperm} I(\max(LRT) \geq \max(LRT^{p})) + 1}{Nperm + 1}$$

where max(LRT) is the observed maximum LRT and $\max(LRT^{p})$ is a permuted maximum LRT. To calculate secondary clusters, this procedure is repeated except only potential clusters that are used are ones that do not overlap with the area of the cluster that was included in the observed maximum LRT. By using the maximum LRT distribution which is more conservative instead of LRT distribution directly holds the type I error for multi-comparisons.

Difference in two methods

Local Moran's I

Local Moran's I is one of the Local Indicators of Spatial Indices (LISA), which provides a local measure of similarity between each region's associated value and those of nearby regions. We can map each location's LISA value to provide insight into the location with comparatively high or low local association with neighboring values. Local Moran's *I* is very similar to Pearson's correlation coefficient, a measure of association between *N* observed values of random variables *X* and *Y*. High values of a LISA suggest clusters of similar (but not necessarily large) counts or proportions across several regions, and low values of a LISA suggest an outlying cluster in a single region (different from most or all of its neighbors) (Lance A Waller and Gotway, 2004).

Spatial scan statistic

A scan statistic involves definition of a moving "window" and a statistical comparison of a measurement (e.g., a count or a rate) within the window to the same sort of measurement outside the window. It considers circular windows with variable radii ranging from the smallest observed distance between a pair of cases to a user-defined upper bound and provides a significance value representing the detected cluster's "unusualness," with an adjustment for multiple testing. At each possible radius in the user-defined interval and for each circle having that radius, we calculate a likelihood ratio statistic testing the constant risk hypothesis versus the specific alternative that risk within regions/locations within the circle differs from the risk in the rest of the study area (Kulldorff, 1997).

The maximum observed likelihood ratio statistic provides an indication of the most likely cluster(s), with significance determined by Monte Carlo testing of the constant risk hypothesis. We generate independent data sets under the null hypothesis, calculate the likelihood ratio statistic for each circle, and store the maximum statistic value, regardless of where it may occur. Statistics are correlated between circles within each simulation, but the maximum values are independent between simulations, providing a valid *p*-value for the most likely cluster, provided that one interprets the *p*-value as the probability of observing a more extreme maximal statistic anywhere in the study area (rather than the significance of observing the maximum at a particular location) (Lance A Waller and Gotway, 2004).

The difference in primary goal, assumption, null hypothesis, search window and detection outputs between local Moran's I and spatial scan statistic are summarized in Table 2-1.

Table 2-1 Comparison of Local Moran's I and spatial scan statistic

	Local Moran' I	Spatial scan statistic
Primary goal	Provide a local measure of similarity between each region's associated value and those of nearby regions.	Determine areas where the observed value appears inconsistent with the value observed over the rest of the study area.
Assumption	First, normality assumption requires that all observations follow identical and independent Gaussian (normal) distributions, but a conditional permutation can be applied; Second, randomization assumption assesses the distribution of the autocorrelation index under random assignment of the values observed to locations	Normality and randomizaion are not necessary. Tests based on scans of local rates often condition on the set of all locations and operationalize the conceptual null hypothesis.
Null hypothesis	Inference for a global index of spatial autocorrelation derives from the null distribution (i.e., the distribution of the index under the null hypothesis). Observed values of the index falling in the tails of this distribution suggest significant spatial autocorrelation.	No clustering through a random labeling or a constant risk hypothesis
Search window	Fixed	Expanding
High/low clusters	May have overlap area between high and low clusters. Current applications have no built-in options to avoid the overlaps	Current applications have built-in option to avoid overlap between high and low clusters
Detection results	The typical output of a LISA analysis involves the values of the LISAs themselves, typically mapped to indicate areas with high values, suggesting stronger local correlation than others within the specified window. Usually detect more clusters than SaTScan. Need to use strict significant level for issue of multiple comparison.	An indication of the location(s) of the most likely cluster(s), often accompanied by some measure of the statistical significance of these cluster(s).No need to worry about multiple comparison

Statistical analyses

Models were first run to assess if, and by how much, individual demographics and SES explained the spatial clustering of BMI. Model 1 used the residuals from a linear regression with BMI in log scale as the outcome, and demographic characteristics (age, race, and gender) as the predictors to find out whether the Local Moran's I and the spatial scan statistic methods identified high or low BMI clusters. The residual for a given person is the model's observed BMI minus the predicted BMI. Given demographic characteristics, individuals with high positive residuals are those with a higher than expected BMI, and individuals with large negative residuals are those with a lower than expected BMI. The same approach was applied to model 2, which further adjusted for education and income.

A second step examined the effect of each one of the seven neighborhood variables on spatial clustering. Linear regressions were used to test differences between the neighborhood features within the high or low BMI clusters and those outside of the clusters. Combinations of seven environmental variables were added to model 2 to test what variables explained the spatial clustering of BMI. A third and final model was run, which added the neighborhood variables with the strongest significance in the tests to model 2.

Maps of high and low BMI clusters were created using ArcGIS (version 10.0). Statistical analyses were conducted using R 2.13.0 (R Foundation for Statistical computing, Vienna, Austria), except for the Local Moran's I method, which used GeoDa (version1.0.1, Spatial

Analysis Laboratory, 2011) and spatial scan method, which used SaTScan (Version 9.1, http://www.satscan.org/).

RESULTS

The mean BMI for the sample was 26.5 (Table 2-2). Most study participants were female, and almost 40% of the sample was 50-64 years old. Higher BMIs were associated with being male, Hispanic, 50-64 years of age, and with lower education, lower income, lower neighborhood property values, and lower residential density.

Table 2-2 Study sample: individual demographic and SES characteristics, neighborhood environment, and BMI

Variable	N	%	Mean BMI (SD)	p-value
Total	1602	100	26.5 (4.9)	
Gender				<0.001
Female	950	59.3	26.0 (5.0)	
Male	652	40.7	27.2 (4.6)	
Age				0.003
18-29	77	4.8	24.8 (4.1)	
30-39	188	11.7	26.3 (5.2)	
40-49	346	21.6	26.2 (4.9)	
50-64	632	39.5	26.9 (5.0)	
65+	359	22.4	26.4 (4.6)	
Race				0.001
White non-Hispanic	1297	81	26.5 (4.9)	
Black non-Hispanic	104	6.5	27.1 (4.3)	
Hispanic	43	2.7	28.6 (5.3)	
Other race	158	9.9	25.6 (4.2)	
Education				<0.001
High school or below	281	17.5	27.4 (5.3)	
Some college/technical school	418	26.1	26.8 (4.9)	
college graduate	903	56.4	26.1 (4.7)	
Household income				<0.001
< $25,000	217	13.5	27.3 (5.7)	
$25,000-34,999	162	10.1	26.5 (5.0)	
$35,000-49,999	254	15.9	26.8 (4.9)	
$50,000-74,999	307	19.2	26.5 (4.8)	
$75,000-99,999	248	15.5	26.7 (4.8)	
> $100,000	414	25.8	25.7 (4.3)	
Neighborhood property value*				<0.001
< $200,000 per residential unit	438	27.3	27.3 (5.1)	
$200,000-300,000 per residential unit	663	41.4	26.7 (5.0)	
> $300,000 per residential unit	501	31.3	25.5 (4.3)	
Neighborhood residential density*				0.142
< 6 residential units per acre	505	31.5	26.8 (5.0)	
6-10 residential units per acre	550	34.3	26.4 (5.0)	
> 10 residential units per acre	547	34.1	26.3 (4.6)	

* Within 833 m of home

In model 1, adjusting for demographic characteristics alone, the Local Moran's I method detected 16 significant high BMI and 44 low BMI clusters (Table 2-3). The spatial scan statistic method yielded one high BMI and one low BMI cluster, both of which overlapped with clusters identified by the Local Moran's I method (**Error! Reference source not found.**, left panel). To visually represent the Moran's I results, the respondents' home neighborhoods (established using a 2-mile radius) in the high or low individual BMI clusters were dissolved if two individual neighborhoods intersected. For model 1, the Local Moran's I clusters dissolved into two "regions" of low BMI clusters in the northern part of the study area, and one region of high BMI clusters in the southern part.

Table 2-3 Number of significant clusters detected by the Local Moran's I and the spatial scan statistic methods

		No. of high BMI clusters	No. of low BMI clusters
Model 1	Local Moran's I	16	44
	Spatial scan statistic	1	1
Model 2	Local Moran's I	9	10
	Spatial scan statistic	1	1
Model 3	Local Moran's I	0	0
	Spatial scan statistic	0	0

Note: Model 1 adjusted for age, gender and race;
Model 2 adjusted for age, gender, race, education and income;
Model 3 adjusted for age, gender, race, education, income, residential density, and property value

In model 2, adjusted for demographic characteristics and individual SES, the Local Moran's I method identified 9 significant high BMI clusters and 10 low BMI clusters. The spatial scan statistic method produced one significant high BMI cluster and one marginally significant low

BMI cluster; again, both of these clusters were located within the clusters identified by the Local Moran's I method (Figure 2-2, right panel). While significant spatial clustering remained, most clusters shrunk in size.

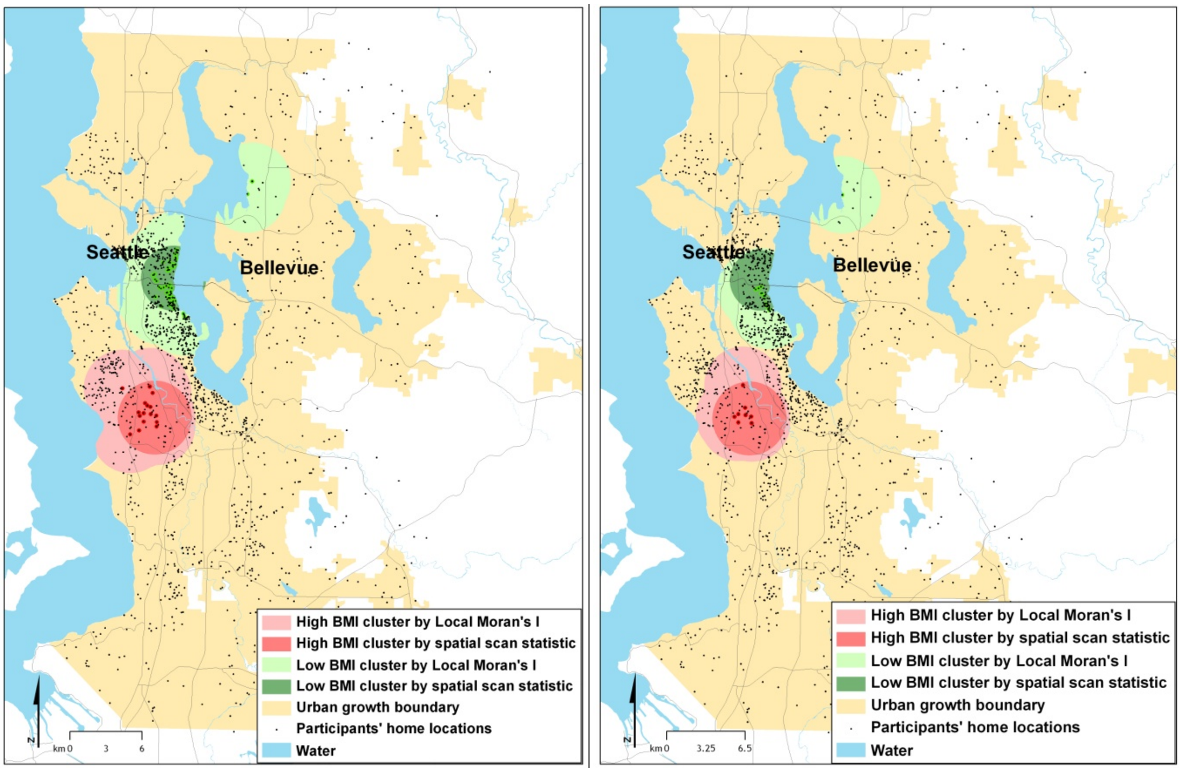

Figure 2-2 Clusters of high and low BMI as detected by the Local Moran's I and the spatial scan statistic methods
Left panel: Clusters adjusted for demographic characteristics (model 1)
Right panel: Clusters adjusted for demographic and SES characteristics (model 2)
Notes:
1) Clusters detected by Local Moran's I are shown in a thin outline and those detected by spatial scan statistic have a bold outline
2) To visually represent the Local Moran's I results, the respondents' home neighborhoods (established using a 2-mile radius) in the high or low individual BMI clusters were dissolved if two individual neighborhoods intersected. The Moran's I clusters identified in models 1 and 2 dissolved into two regions of low BMI clusters (simple hatch) in the northern part of the study area, and one region of high BMI clusters (crosshatch) in the southern part.
3) The grey area represents the urban growth boundary, which contains the built-up areas of King County

The neighborhood environment of areas within high and low BMI clusters differed from that of areas not assigned to either type of cluster (Table 2-4). Residential density and property values were significantly lower in the high BMI clusters and significantly higher in the low BMI clusters than in areas outside either cluster. Low BMI clusters had a significantly higher density of supermarkets, grocery stores, and street intersections compared to areas outside the clusters; they also had significantly more area in parks. High BMI clusters had a significantly lower density of fast food and quick service restaurants than areas outside the clusters. All other relationships were not significant.

Table 2-4 Neighborhood environment features; mean value within 833 m of home, and with adjustment for individual demographic and SES

	Not in clusters (REF)	In high BMI cluster	In low BMI cluster
N	1366	60	176
Residential density units (unit per acre)	10.2	7.00**	14.70***
Property value per residential unit ($ 1,000)	275.11	190.88***	321.42***
Density of supermarkets and grocery stores (count)	1.56	0.95	2.87***
Density of fast food and quick service restaurants (count)	13.41	2.98*	15.36
Percent of park land (% of area)	6.03	5.04	7.10*
Intersection density (count)	154.59	137.48	220.36***
Density of traffic volume (annual average daily traffic per km of street)	58,356	45,493	71,031

*P<0.05, **P<0.01, ***P<0.001

The two neighborhood environment variables that were significantly associated with both high and low BMI clusters, property values and residential density, were added in model 3, which was also adjusted for demographic characteristics and individual SES. In this model, no cluster was identified by either detection method.

DISCUSSION

Obesity was found to concentrate spatially in high and low BMI clusters using two detection techniques in conjunction with the geocoded home addresses of individual obese and non-obese cases. Seattle King County presented a North – South obesity gradient, with the northern part being thinner, richer, better educated and more densely populated, and the southern part being more obese, lower income, and less densely populated. As for the influence of the respondents’ home neighborhood environment, two characteristics, residential density and property values, were consistently related to high and low BMI clusters, while such built environment features as the density of food sources, streets, and parks were not. In fact, residential density and property values appeared to completely explain geographic disparities in obesity as no cluster was predicted after these two attributes were added to the model adjusted for sociodemographic variables.

Of note, property values explained clustering and BMI variation across space more strongly than did residential density: the adjusted model with only property values predicted fewer clusters than that with only residential density (only 3 high BMI and no low BMI clusters for the Local Moran's I model with property values, versus 9 high and 3 low BMI clusters for the model with residential density, data not shown). In contrast, residential density has previously been associated with health-related behaviors (Ewing et al., 2008; Frank et al., 2005; Kligerman et al., 2007). The link found in this study between higher residential density and low BMI might at first seem counter-intuitive, as populations in the higher SES categories often live in lower density neighborhoods. However, the mean densities of low BMI clusters in the present sample were relatively low, consistent with dense single-family housing, while the lower densities found to be associated, albeit more weakly, with high BMI clusters corresponded to that of sprawling suburban-style subdivisions.

Finding that obesity clusters are effectively explained by property values suggested that the geographic distribution of obesity might be driven by neighborhood economic factors. The assumption, which should be tested further, was that residential property values reflected certain neighborhood features thought to be either obesogenic or protective of obesity. Specifically, proximity to supermarkets, parks, and opportunities for physical activity might also be associated with higher property values. It was previously found that residential property values were a powerful predictor of women's body weight: women living in homes with the lowest property values were 3.4 times more likely to be obese than women living in homes with the highest property values, after adjusting for education and income (Rehm et al., 2012). These findings suggested that neighborhood property values might serve to identify the locations of prevention and intervention programs.

Spatial analysis primarily serve to detect patterns of events (L A Waller and Gotway, 2004). Importantly, the two cluster detection techniques used in this study yielded convincingly similar results while relying on different assumptions and mechanisms to identify spatial clustering. Future research should also evaluate the effectiveness of some of the distinctive techniques introduced by this study. Specifically, the individual-level analyses represented a departure from research using different levels of geographic aggregation. Also important was the use of residuals from regression models was used as the outcome of the analyses in order to specify error dependence structure for inference. Running regression models to estimate unmeasured neighborhood-level factors from residuals on the spatial distribution of obesity further helped identify areas with clustering of higher or lower than expected BMI. Finally, the models took into account differences in the built environment features between individual neighborhoods that were inside and those outside of the clusters.

This study had limitations. While the significant association found between neighborhood features and BMI clusters suggested potential neighborhood influences on BMI, it did not imply causality or generalizability. Another cross-sectional study of 1,863 adults living in Metro Vancouver, BC, did not observe associations between obesity clusters detected with the Local Moran's I and neighborhood residential density or income (Schuurman et al., 2009). However, the study's population was different from that of the present study: it came from 8 discrete suburban neighborhoods and had relatively low rates of obesity (16% versus more than 21% in this study). At the methodological level, the distance thresholds selected for the spatial weights in the Local Moran's I and the maximum cluster size in spatial scan statistic were hypothesis-driven and would need to be tested. Finally, the circular buffers used to calculate the neighborhood variables might not represent actual neighborhood boundaries, which typically have complex shapes (Kulldorff et al., 2006; Patil and Taillie, 2004; Kunihiko Takahashi et al., 2008).

CONCLUSION

This study's spatial cluster analyses examined geographic variation in BMI in relation to neighborhood characteristics. The Local Moran's I and the spatial scan statistic methods identified a similar pattern of high and low BMI clusters, which suggested that the geographic distribution of BMI was not random and might be related to features of the local neighborhood. Further examination of differences in neighborhood features within and outside BMI clusters showed that property values were key to explaining clustering. This variable needs to be considered in future research on the built environment and health in order to help target the location of prevention and intervention strategies.

Chapter 3 : MEASURING NEIGHBORHOOD WALKING AND WALKABILITY

ABSTRACT

Understanding how neighborhood characteristics correlate with each other and how they influence walking is a priority in active living research. The more frequently people walk around their neighborhood, the more likely they are exposed the same built environment and less likely they remain sedentary for a long period of time. Walking and non-walking physical activity bouts were identified via integrated accelerometer, portable GPS, and 7-day travel log data from 675 TRAC (Travel Assessment and Community) study participants at baseline. Home addresses were geocoded and neighborhood walking bouts were defined as those occurring within 0.5 mile buffer around each participant's home. Participant-specific spatially continuous objective values of neighborhood environment features (SmartMaps) were constructed for seven environment variables. Collinearity among neighborhood environment variables was analyzed and variables that were strongly correlated with residential density were excluded in the regression analysis to avoid erroneous estimates. A Zero Inflated Negative Binomial (ZINB) served to estimate associations between home neighborhood environment characteristics and neighborhood walking frequency, adjusting for socio-demographic characteristics. The study found that more than half of participants' walking bouts occurred in their own home neighborhood and revealed that residential density or job density needed to increase substantially to effectively support walking behavior in home neighborhood. Finding that a large proportion of walking takes place in the home neighborhood highlights the importance of continuing to examine the impact of the home neighborhood environment walking. Potential interventions to increase walking behavior may benefit from changing the built environment of residential areas. People in the present study living in neighborhoods with higher residential and job densities walk more frequently in their neighborhood.

INTRODUCTION

Physical activity (PA) is defined as any body movement produced by skeletal muscles that requires energy expenditure. Insufficient physical activity (PA) is the fourth leading risk of global mortality causing an estimated 3.2 million deaths globally (World Health Organization (WHO), 2009). Regular, moderate - intensity physical activity (e.g. brisk walking, dancing or household chores) can reduce the risk of cardiovascular diseases, diabetes, colon and breast cancer, depression, and other chronic disorders (World Health Organization (WHO), 2012). It is recommended that adults aged 18-64 years engage in at least 150 minutes of moderate-intensity aerobic PA throughout the week (World Health Organization (WHO), 2010).

Among moderate-intensity PA, walking is a popular, convenient, and free form of PA that can be incorporated into everyday life and sustained into old age (Morris and Hardman, 1997; Ogilvie et al., 2007; Siegel et al., 1995). These characteristics have made the promotion of walking a promising public health strategy to counter physical inactivity trends (Eyler et al., 2003). However, it is estimated that 60-80% of the world's population does not meet the recommendations (World Health Organization (WHO), 2007). In a world dominated by cars or other forms of motorized transport, walking is a location-dependent activity that concentrates in relatively small, neighborhood-sized areas (Rappaport and Seidman, 2000). Accordingly, research on the contribution of walking to physical activity has examined neighborhood-level determinants of walking, and specifically the effects of the neighborhood environment on walking (Brownson et al., 2005; Kawachi and Berkman, 2003). To date, most of this research has focused on the characteristics of the residential neighborhood, based on the hypothesis that people's place of residence exerts the most influence on walking behavior. Yet little is known about where people actually walk. In fact, since most people spend a substantial amount of time outside their home neighborhood, it is conceivable that they are exposed to a range of environments other than that of their home neighborhood, which may influence their walking behaviors.

Most studies seeking to find out where people walk have been limited by their use of self-reported data on the behavior. Self-report data are not only imprecise, e.g. the International Physical Activity Questionnaire (IPAQ) typically over-estimates walking frequency (Lee et al., 2011; Rzewnicki et al., 2003) while transportation surveys focusing on motorized trips tend to underreport walking trips (Stopher and Greaves, 2007), but many also do not provide information about the location where walking occurs. And if the data collected refer specifically to walking in the home neighborhood, then inter- and intra-participant variations in the definition of neighborhood (e.g., different people considering size of their neighborhood to be different) may bias results.

Furthermore, there are measurements issues related to capturing the characteristics of the neighborhood built environment. Many studies use composite indices of neighborhood environment such as walkability index (Norman et al., 2013) and sprawl index(Ewing et al., 2008) to combine measures of many aspects of built environment, reduce measurement error, collinearity and over-adjustment. However they introduce methodological concerns such as validity, reliability and generalizability, as such indices are typically developed for a specific setting. Hence, simple measures of individually observable variables are considered to be more, or at least equally, effective in characterizing environments for walkability than composite indices (Lee and Moudon, 2006a). The controversy is found in a study that observed significant association between sprawl index and obesity, but not between its components and obesity

(Kelly-Schwartz et al., 2004). Another drawback of a composite index is that it is less useful for intervention because one cannot identify specific components that have highest priory to change, and it is less informative for evaluating specific hypothesizes (Feng et al., 2010). Other studies find significant association between individual neighborhood characteristics and walking, but ignore the fact that neighborhood environment variables are highly correlated. The use of composite indices may lead to erroneous estimates of coefficients, large standard errors in the related independent variables and spurious significant results due to multicollinearity, though it does not reduce the predictive power of the model as a whole.

This study fills a gap by using objective data on the frequency, duration, and location of walking assessed over the course of 7 days. These data preview habitual walking as it occurs inside or outside of the home neighborhood, and therefore open the opportunity to test whether the characteristics of the home neighborhood environment are indeed associated with walking.

METHODS

Participants

Participants in the present analysis were the baseline sample of the Travel Assessment and Community (TRAC) study examining the impact of a new light rail system on PA. Participants were selected to reside proximal (case) or distal (control) from future light rail stops, but to be living within the same county (Seattle/King County, WA) and living in areas with similar built environments (defined by residential density, housing type, home values, bus transit access, and availability of proximate neighborhood services) and census-based demographic characteristics (household income and race/ethnicity). Eligible households in identified areas were contacted using address and phone information from marketing companies. Eligible participants needed to be 1) 20+ years old, 2) able to complete the travel log and survey in English, and 3) able to walk unassisted for at least 10 minutes. Participants consented to participate and the study was approved by the Seattle Children's Institutional Review Board. At least one valid assessment day was available for 701 participants. Of these, 18 were excluded because they did not have at least one valid day with GPS coverage totaling three minutes; eight more were excluded because they did not complete the attitudinal/demographic survey. The remaining 675 participants contributed 4494 valid person-days of observation.

Data collection

Eligible and interested participants were mailed an accelerometer (Actigraph GT1M), portable GPS device (GlobalSat DG-100), and 7-day paper travel log. Participants were also provided a written or on-line (based on their preference) attitudinal and demographic survey to complete. Soon after receiving these materials, participants were contacted by study staff to review procedures (e.g., how to wear the devices; how to charge the GPS device nightly) and asked to wear the accelerometer and GPS for seven days during waking hours and to complete the travel log for these days. Accelerometer data were aggregated to 30-second epochs and GPS devices were set to collect data at 30-second intervals. Participants mailed backed the devices and travel log (and survey if in written form) in a pre-paid envelope (Saelens et al., 2014).

Socio-demographic data on participants' age, sex, household income, race/ethnicity, highest level of education came from the survey as well as height and weight from which BMI was calculated.

Data processing

Physical activity was divided into walking and non-walking activity bouts following a process by which accelerometer data were integrated with GPS and travel log information. The process is described elsewhere (Kang et al., 2013). Briefly, bouts of ≥5 minutes of accelerometer counts averaging >500 per 30-second epoch were considered to be physical activity; these bouts were then considered to be walking based on GPS speeds and/or on temporal overlap or proximity to walking trips recorded in the travel log. Travel logs included the places and durations reported to be visited and the travel mode between places.

In the present analysis a day was considered valid if it had at least one place record in the travel diary and an accelerometer wearing time of ≥ 8 hours per day. Accelerometer periods of ≥ 20 minutes with continuous zero values were considered nonwearing times. An assessment day may or may not have had GPS data. The final sample for the present analysis consisted of 675 participants and 4,494 person-days (mean=6.7 days/person; SD=1.7).

Outcomes and neighborhood environment

A walking bout was defined as a non-mechanical and human-powered travel associated with sustained light and moderate intensity physical activity for at least 7 minutes with a 2 minutes tolerance of lower PA intensity. The definition served to isolate “walking as travel in space” from other types of PA (Kang et al., 2013). Operationally, walking bout lines were walking bouts with at least two GPS points.

Home neighborhood was defined as 0.5 mile (833 meters, 10 minutes walking distance) radius around home location, which was geocoded using ArcGIS 10.0. Figure 3-1 shows the geographic distribution of home locations and walking bouts. In order to study home neighborhood walking, walking bouts with at least one GPS point falling within 0.5 mile buffer around home location were selected and considered as walking in the home neighborhood.

The outcome of the analysis was the frequency of daily walking in home neighborhood, which was measured by dividing the total number of that participant's assessment days, representing habitual walking around home. The more frequently the people walked around the neighborhood, the more likely the participant was exposed the same built environment and less likely the participant remained sedentary for a long period of time. Home neighborhood daily walking frequency was highly correlated with home neighborhood daily walking duration in the TRAC study sample with correlation 0.87, which indicated that the participants who walked frequently also have longer walking duration per day. Thus, home neighborhood daily walking bout frequency, which also represents home neighborhood daily walking bout duration in the study, is used in the subsequent regression models.

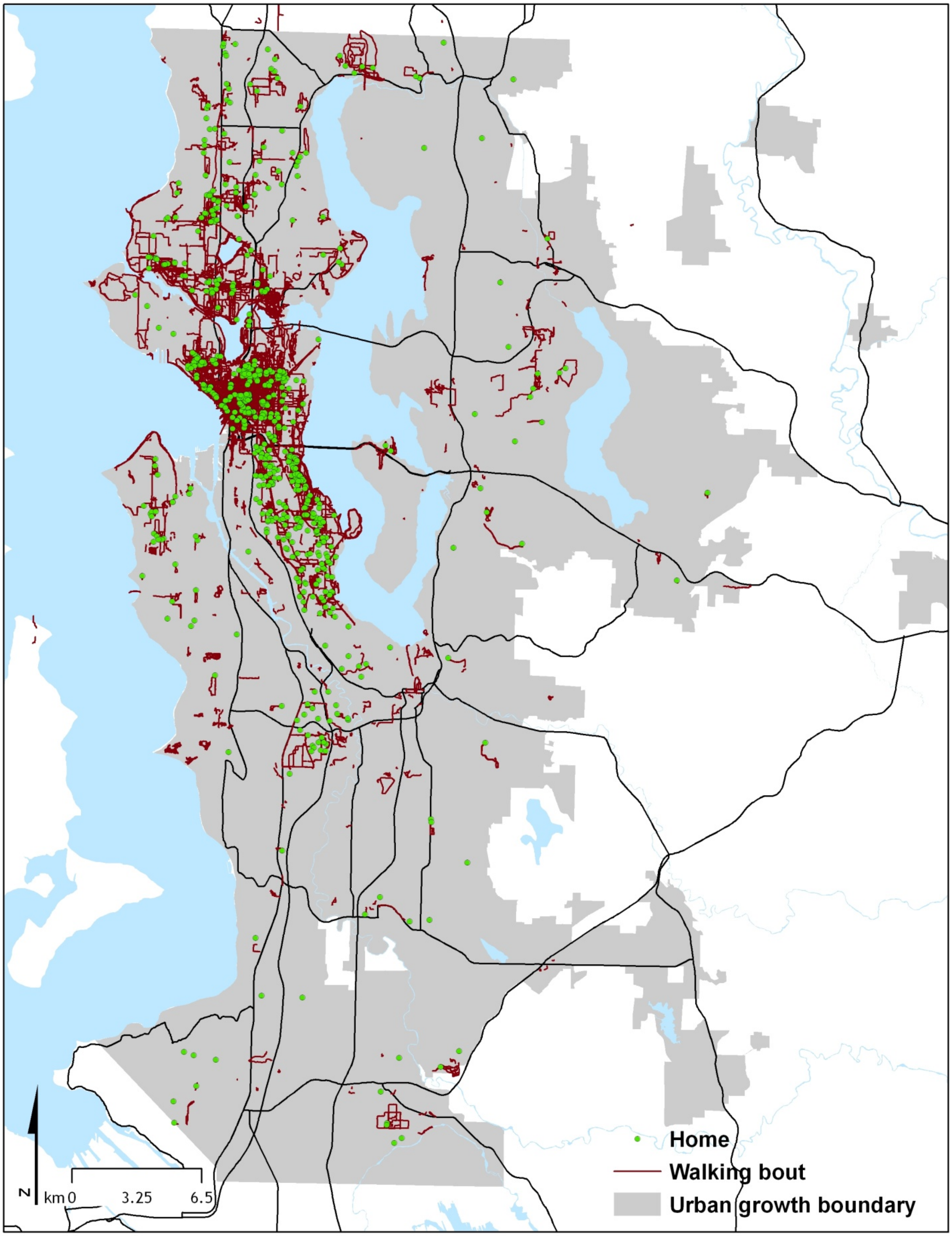

Figure 3-1 The geographic distribution of participants' homes and walking bouts

Neighborhood data and measures

Seven neighborhood environment variables were selected in the domains of neighborhood composition, environment support for active living, and transportation, because they had previously been associated with physical activity. Residential density and job density were measures of neighborhood composition (Ewing et al., 2008; Feng et al., 2010; Forsyth et al., 2007; Papas et al., 2007); and residential property values represented neighborhood socioeconomic status (SES) (Moudon et al., 2011a; Rehm et al., 2012). Amenities specific to physical activity include included the number of parks in the neighborhood (Feng et al., 2010; Papas et al., 2007) and fitness density (Moore et al., 2013). The transportation environment variable included street intersection density and sidewalk density (Block et al., 2011; Feng et al., 2010).

Neighborhood environment values came from SmartMaps, which were created using focal raster processing in GIS (Hurvitz and Moudon, 2012). First, the study area parcel characteristics were translated into a 30x30 m grid, previously shown to represent urban and suburban parcels with sufficient spatial fidelity (Moudon et al., 2011b). The neighborhood around each grid cell was defined within an 833 m radius (a 10-minute walking distance). The mean value of the environmental variables was calculated for each grid cell using ArcGIS (version 10.0) Spatial Analyst Extension. Environmental variables were thus converted to a continuous function at each grid point, and then linked spatially to our study participants' geocoded home addresses for analyses.

Data analysis

Summary statistics were calculated for duration, frequency of walking activity and seven neighborhood environment variables in home neighborhood of the study population. Means and standard deviations compared neighborhood environment variables between participants with no walking bouts at all and those with at least one walking bout, and among the latter group, those with at least one walking bout in home neighborhood and those without during assessment period.

Multicollinearity in the seven home neighborhood environment variables was examined with a correlation matrix to avoid inaccurate estimates of coefficients and standard errors. Variables with correlation coefficient greater than 0.7 were excluded from the regression models. Variable exclusion criterion is based on measurement accuracy and significance as reported in the literature. Conditional indexes and variance decomposition proportions (Belsley, 1991), regression collinearilty diagnostic procedures, were also implemented to confirm the exclusion by using correlation coefficient.

Two regression models estimated the frequency of walking in the home neighborhood. First, a zero-inflated negative binomial (ZINB) regression was run for the entire population, which included participants with no walking bout during the assessment period. This model takes into account excessive number of zeros and overdispersed distribution in the outcome variable. A likelihood ratio test for over-dispersion in count data was used to compare the log-likelihoods of a negative binomial regression model and a Poisson regression model (Cameron Trivedi, P. K., 1998). A Vuong's non-nested hypothesis test was also used to compare zero-inflated count models with their non-zero-inflated analogs, i.e. zero-inflated negative-binomial versus ordinary negative-binomial (Vuong, 1989). Both tests justified the use of negative zero-inflated negative binomial regression to model daily walking bout frequency for the entire study population. Second, a negative binomial regression was carried out for the population who had at least one walking bout during assessment period, excluding excessive zeros (those who did not have any walking bouts during assessment period).

RESULTS

The majority of the sample was female, between the ages of 40 and 65, non-Hispanic White, with a college education, and of normal weight. Nearly 90% of the participants had at least one walking bout and more than 10% had no walking bout at all during the assessment period.

As shown in the density plot of GPS coverage of walking bouts (GPS coverage ratio

Figure 3-2), ten percent of walk bouts have a GPS coverage of less than 50%. In terms of walking frequency, the total number of walking bouts in the home neighborhood for the study population is 3192 and the total number of walking bouts is 5628, thus 57% of walking occurred in home neighborhood. In terms of duration, 63% of walking occurred in home neighborhoods. On average, the daily count of walking bouts per person in the home neighborhood was 0.7 (compared to a total number of daily walking bouts of 1.2) and the daily walking duration per person in home neighborhood was 12 minutes across all assessment days (compared to 18.5 total daily minutes).

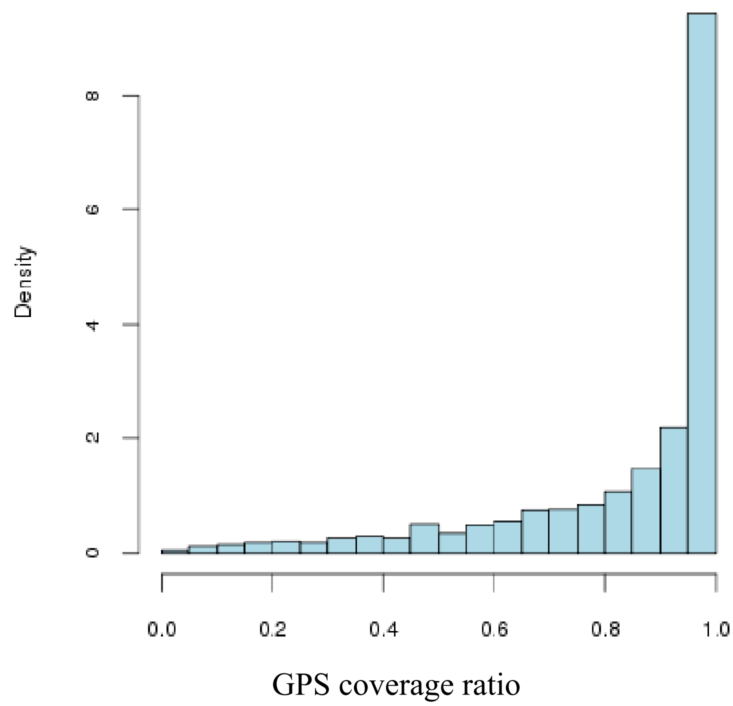

Figure 3-2 walking bout GPS coverage ratio

Comparing home neighborhood attributes, all neighborhood environment variables were significantly higher among those participants who had at least one walking bout in the home neighborhood than those who had none (Table 3-1). However, further correlations between neighborhood environment variables lead to the exclusion of street intersection density, sidewalk density, and fitness facility density from the subsequent regression analyses to avoid multicollinearity (correlation matrix of Table 3-2). These variables can be linearly predicted from residential density.

Table 3-1 Descriptive characteristics of study sample

	All (n=675)		Population With > 0 walking bout anywhere (n = 605)		Population with >0 walking bout in neighborhood (n=497)		Population with no home neighborhood walking bout (n = 108)		
	Mean / Percent	SD	Mean / Percent	SD	Mean / Percent	SD	Mean / Percent	SD	p-value
Gender									
Male	36.9%		38.8%		38.4%		32.8%		
Female	63.1%		61.2%		61.6%		67.2%		
Age									
18-39	21.7%		22.0%		23.6%		16.4%		
40-65	63.5%		64.9%		62.4%		66.7%		
>65	14.8%		13.1%		14.0%		17.0%		
Race									
Hispanic or non-white	20.4%		19.2%		18.5%		25.7%		
Non-Hispanic White	79.6%		80.8%		81.5%		74.3%		
Education									
Less than college graduate	29.6%		27.4%		26.3%		39.1%		
College graduate	70.4%		72.6%		73.7%		60.9%		
Weight status									
BMI < 25	48.5%		50.0%		52.9%		35.6%		
BMI = 25-30	30.1%		29.7%		28.8%		33.8%		
BMI >30	21.4%		20.3%		18.3%		30.6%		
Home neighborhood walking Bouts									
Daily walking bout count	0.7	(0.8)	0.8	(0.8)	1.0	(0.8)	0.0	(0.0)	
Daily walking bout duration	12.0	(15.3)	13.4	(15.6)	16.3	(15.8)	0.0	(0.0)	
Neighborhood environment									
Residential density	20.1	(14.3)	20.7	(14.4)	22.2	(14.9)	14.3	(10.1)	<0.001
Residential property value	247.3	(81.8)	249.8	(83.7)	251.2	(84.1)	236.3	(74.2)	0.027
Job density	32.8	(61.8)	34.3	(63.1)	38.9	(66.5)	15.9	(41.7)	<0.001
Street intersection density	214.0	(74.3)	216.6	(74.4)	221.9	(75.6)	191.8	(65.9)	<0.001
Sidewalk length	150.8	(80.7)	154.7	(80.0)	161.4	(78.7)	121.1	(79.0)	<0.001
Fitness	5.5	(5.0)	5.7	(5.0)	6.2	(5.1)	3.7	(4.0)	<0.001
Park count	6.2	(3.6)	6.4	(3.6)	6.6	(3.6)	5.3	(3.6)	<0.001

Table 3-2 Correlation matrix of BE variables

	Residential density	Residential property value	Job density	Street intersection density	Sidewalk length	Fitness count	Park count
Residential density	1.00	-0.28	0.62	0.78	0.75	0.80	0.38
Residential property value	-0.28	1.00	-0.13	-0.13	-0.02	-0.23	0.14
Job density	0.62	-0.13	1.00	0.73	0.45	0.77	0.27
Street intersection density	0.78	-0.13	0.73	1.00	0.84	0.77	0.48
Sidewalk length	0.75	-0.02	0.45	0.84	1.00	0.70	0.46
Fitness count	0.80	-0.23	0.77	0.77	0.70	1.00	0.32
Park	0.38	0.14	0.27	0.48	0.46	0.32	1.00

The neighborhood environment covariates remaining in the models, residential density, residential property value, job density, and the number of parks, were classified into tertile ranges to aid in the interpretation of model results and to take into account their possible non-linear relationship with the outcomes (Table 3-3).

Table 3-3 Tertile ranges of neighborhood environment variables

	Low tertile	Medium Tertile	High tertile
Residential density (unit/acre)	[2.45, 10.62)	[10.62, 22.44)	[22.44, 62.63)
Residential property value ($1000)	[100, 205)	[205, 259)	[259, 873)
Job density (jobs/acre)	[0, 3.7)	[3.7, 12.4)	[12.4, 272.3)
Park (count)	[0, 4)	[4, 8)	[8, 19)

The negative binomial part in the ZINB for the entire study population yielded four variables, gender, weight status, residential and job density, which were significantly associated with having more walking bouts in the home neighborhood (Table 3-4). For male, the expected count of walking bouts in home neighborhood was 23% higher than female. For the obese, the expected count of walking in home neighborhood was 45% lower than those of normal weight. The expected count of walking bouts around home for those living in the high tertile residential density neighborhood was 43% higher than those living in the low tertile residential density neighborhood. Finally, the expected count of walking bouts around home for those living in the high tertile job density neighborhood was 60% higher than those living in the low tertile job density neighborhood. Note that the binomial part of the model is used to explain the excessive zeros in the outcome, narrowing down the confidence interval in the negative binomial part.

Table 3-4 ZINB results (Outcome: # of walking bouts per day around home)

	exp(β)	Confidence interval
Negbin with log link		
Male *	1.234	(1.027 , 1.483)
Age40-65	1.130	(0.915 , 1.396)
Age>65	1.235	(0.913 , 1.671)
Non-Hispanic white	0.839	(0.660 , 1.066)
Income50-100K	0.975	(0.788 , 1.206)
Income>100k	0.910	(0.698 , 1.186)
College graduate	1.058	(0.853 , 1.311)
Weight status-Overweight	0.865	(0.695 , 1.078)
Weight status-Obese ***	0.549	(0.426 , 0.708)
Residential density-Medium	1.219	(0.929 , 1.599)
Residential density-High*	1.434	(1.003 , 2.049)
Property value-Medium	1.120	(0.874 , 1.435)
Property value-High	0.921	(0.692 , 1.226)
Jobs density-Medium	1.182	(0.910 , 1.536)
Jobs density-High *	1.616	(1.102 , 2.370)
Park count-Medium	0.894	(0.705 , 1.132)
Park count-High	0.844	(0.639 , 1.116)
Binomial with logit link		
Male	0.708	(0.291 , 1.726)
Age40-65	2.267	(0.399 , 12.871)
Age>65	4.509	(0.641 , 31.725)
Non-Hispanic white	0.565	(0.240 , 1.329)
Income50-100K	1.040	(0.441 , 2.451)
Income>100k	0.684	(0.182 , 2.578)
College graduate	0.684	(0.308 , 1.516)
Weight status-Overweight	1.776	(0.630 , 5.003)
Weight status-Obese	2.301	(0.806 , 6.564)
Residential density-Medium	0.909	(0.343 , 2.408)
Residential density-High	0.145	(0.012 , 1.722)
Property value-Medium	0.918	(0.374 , 2.255)
Property value-High	0.353	(0.075 , 1.660)
Jobs density-Medium	1.109	(0.412 , 2.987)
Jobs density-High	0.787	(0.130 , 4.779)
Park count-Medium	0.773	(0.313 , 1.908)
Park count-High	1.277	(0.331 , 4.926)

Note: *p value <0.05, ** p value<0.01, *** p value< 0.001

For the subpopulation who had at least one walking bout during the assessment period, a negative binomial regression (Table 3-5) was used to test whether it yielded similar results to those presented for the entire population. Gender was no longer significantly related to more frequent walking in the neighborhood, but the odd of having more neighborhood walking bouts was similarly lower for the obese. Also, the effect of residential density and job density on walking frequency was stronger for this subpopulation.

Table 3-5 Negative binomial regression for participants who had at least one walking bouts

	exp(β)	**Confidence interval**
Male	1.190	(0.994 , 1.423)
Age40-65	1.009	(0.817 , 1.245)
Age>65	1.135	(0.842 , 1.529)
Non-hispanic white	0.887	(0.707 , 1.114)
Income50-100K	0.955	(0.778 , 1.173)
Income>100k	0.882	(0.691 , 1.127)
College graduate	1.052	(0.855 , 1.293)
Weight status-Overweight	0.842	(0.688 , 1.031)
Weight status-Obese ***	0.529	(0.416 , 0.673)
Residential density-Medium	1.225	(0.975 , 1.537)
Residential density-High**	1.606	(1.141 , 2.261)
Property value-Medium	1.155	(0.921 , 1.448)
Property value-High	1.040	(0.824 , 1.312)
Jobs density-Medium	1.186	(0.938 , 1.499)
Jobs density-Hight **	1.678	(1.165 , 2.416)
Park count-Medium	0.916	(0.728 , 1.153)
Park count-High	0.828	(0.634 , 1.082)

Note: *p value <0.05, ** p value<0.01, *** p value< 0.001

In addition, utilitarian and recreation walking in home neighborhood were separately fitted using the models above (Not shown). For utilitarian walking, both medium and high residential density and job density tertiles were significantly positively associated with walking bouts frequency. However, none of the home neighborhood environment attributes had a significant relationship with recreational walking.

DISCUSSION

Using objective measures of walking frequency and location, this study confirms the value of the many past studies focusing on the influence of the home neighborhood environment on walking activity. At the population level, more than half of the walking bouts occurred in the home neighborhood. Of the participants who had at least one walking bout, 82% had at least one walking bout in their neighborhood. Also, people who walked in their neighborhoods walked more frequently than those who did not: they had an average of 1.6 walking bouts per day versus 0.6 for those who did not walk in their neighborhood. At the individual level, the frequency of walking in the neighborhood was associated with being male and having a lower BMI, but with no other sociodemographic characteristic.

The study supports the growing evidence that walking in the home neighborhood is significantly more frequent in areas of higher residential and employment density. However, the results suggest that residential density or employment density need to increase to higher levels to actually support higher levels of walking. The lowest bound of highest tertile of residential density (gross density of 22 units/acre) is equivalent to a neighborhood where row houses are dominant. Finding that high job density is associated with more walking in a residential neighborhood implies that activity-supporting neighborhoods intermix residential and employment uses. Areas combining residential and employment uses may also encourage walking because of their higher level of transit service, higher parking costs, and correspondingly lower automobile ownership rates (Cervero, 2006; Frank and Pivo, 1994).

The study clarifies that residential and employment densities are the primary aspects of the home neighborhood environment related to more walking. Past research found that higher residential density, greater street connectivity, greater number and variety of destinations, and mixed land use in residential neighborhoods were associated with more walking (Coogan et al., 2009; Duncan et al., 2010; McConville et al., 2011; Saelens and Handy, 2008). The present study demonstrated that residential density was highly correlated with street intersection and sidewalk density, as well as with the density of fitness facilities. These results suggest that measuring residential and employment density may be sufficient to assess the walkability of a neighborhood. Furthermore, analyses adding other neighborhood characteristics to these two variables likely engender issues of multicollinearity, which can lead to erroneous estimates of coefficients, large standard errors, and potentially spurious significant results. There is a caveat to these findings: in this study, residential and employment densities were calculated using fine-grained, parcel-level data, which may not obtainable in all settings. Studies based on coarser census data have used composite indices of the neighborhood environment (e.g., the walkability index (Norman et al., 2013) and the sprawl index(Ewing et al., 2008) to circumvents the issue of multicollinearity. Yet such indices raise concerns about their validity, reliability and generalizability, as they are developed within a specific setting and thus only applicable under certain context. One study observed significant associations between the sprawl index and obesity, but not between the index components and obesity (Kelly-Schwartz et al., 2004). Furthermore, composite indices cannot be used to design policy interventions or to test specific hypothesizes (Feng et al., 2010) because of the inability to identify specific attributes of environment that are modifiable and have the highest potential to change behaviors. Hence, if fine-grained data are available, it is preferable to use such simple measures of individually observable variables as residential and employment density (Lee and Moudon, 2006a). This study shows that these two variables are effective in characterizing environments for walkability.

Finally, the study supports previous findings that utilitarian walking is more sensitive to increase in residential density and job density, but recreational walking is not related to these two factors (Lee and Moudon, 2006b).

The study has limitations. First, the sample was drawn from only one U.S. metropolitan region, restricting the generalizability of findings. This study was part of a large study investigating the impact of light rail on travel behavior and physical activity, and participants had more transit access than the general local or U.S. population. While walking in the neighborhood was objectively measured, missing GPS data may have resulted in inexact estimates of home neighborhood walking bouts.

CONCLUSION

Objective measures of walking show that walking takes place most frequently in the home neighborhood. Higher densities of both residential and employment uses effectively support walking behavior in home neighborhoods. These densities are correlated with high street intersection and sidewalk density. The study confirms the importance of past research focusing on the walkability of the home neighborhood and indicates that efforts to increase development density in residential neighborhoods will help promote more walking.

Chapter 4 : QUANTIFY THE EFFECT OF LIGHT RAIL ON WALKING AROUND STATIONS: A LONGITUDINAL STUDY

ABSTRACT

Light Rail Transit (LRT) serves a high number of transit users and LRT stations can encourage walking and cycling and hence improve quality of life, health and environmental sustainability. The environment around LRT stations can encourage residents to not only walk between their residencesand transit, but also to stores and other activities near the stations. A total 230 participants living around 13 LRT stations were selected from the baseline (2008) and post 1 (2010) sample of the Travel Assessment and Community (TRAC) study to assess changes in walking around LRT stations after the completion of LRT. Walking and non-walking physical activity bouts were identified by an algorithm using accelerometer, portable GPS, and 7-day travel log data. Home addresses were geocoded and walking bouts around LRT stations were identified with quarter mile buffers around each LRT station. The proportion LRT station walking was calculated as daily LRT station walking divided by the daily overall walking of each participant, representing the spatial concentration of walking around LRT stations for each participant at baseline and post 1. Proportion of LRT station walking at baseline was positively associated with that at pose 1. In addition, participants who were college graduates and frequent light rail users significantly increased their proportion of LRT station walking; overweight and obese participants significantly decreased their proportion of LRT walking. The distance from the residential location to the nearest light rail station was negatively related to the change of proportion of LRT station walking bouts. The shift of individual walking behavior into LRT station areas is associated with weight status, distance from home to the nearest station and light rail usage frequency.

INTRODUCTION

Light Rail Transit (LRT) plays an important role in Smart Growth, an approach to urban form that seeks to counteract sprawl by concentrating growth in compact, walkable and transit supportive centers in order to improve quality of life, health and environmental sustainability. Following Smart Growth, urban forms encourage walking, cycling and improves access to transit (Frank et al., 2006). Smart Growth planning calls for LRT station areas to become neighborhoods supporting the needs of residents or workers beyond those of transit riders. These transit-related neighborhoods have long been promoted as Transit-Oriented Development (TOD), characterized by relatively dense, mixed use, and pedestrian friendly environments (Calthorpe, 1993; Cervero, 2001). Thus LRT systems can create urban corridors where each station area can become a node of activity where walking, bicycling, and transit use are the most convenient forms of travel. LRT station areas are often re-designed with traffic-calming measures (e.g., narrow streets, stop signs, special street treatments, and traffic diverters) to further favor pedestrian mobility.

LRT systems with TOD urban forms at each station can have a considerable impact on mobility and accessibility. Because high-performing LRTs have the majority of their riders access transit by walking, TOD-like station areas are more likely to generate transit riders than lower density, automobile-oriented areas. A California study found that for every additional 100 employees per acre near a station area, rail ridership rose 2.2 percent; walking as the access mode from home to rail station and from rail station to workplace is 87.8% and 74.2% respectively of all travel modes (Cervero, 1994). Also, changing station-area built environment to make it more TOD-like can lead to more residents or workers being willing and able to substitute automobile trips with transit. In addition, environmental improvements around LRT stations can encourage people to walk not only between their residences or offices and the light rail station, but also to local stores and other routine daily destinations. Improved streetscapes, street trees, lighting, walkways, reduced street crossing distances, crossing signal timing, and other facilities can further help create an environment that is more conducive to walking.

Light rail systems have been shown to support health-related behaviors. A longitudinal study revealed that walking to light rail was positively associated with moderate physical activity (Brown and Werner, 2007); light rail riders had more moderate physical activity bouts, greater neighborhood satisfaction and fewer car rides than non light rail riders (Brown and Werner, 2008); station area residents experienced greater than anticipated perception of safety while walking, as well as feeling that their children were safer; finally, there was increased neighborhood satisfaction after the operation of light rail (Brown and Werner, 2011).

The effect of distance from either home or work on transit access and especially on walking to transit has long been examined. However, changes in amounts of walking around stations areas after the implementation of new or improved transit service has not been fully investigated. While longitudinal data have been collected for walking to transit, little is known about how individuals may change their walking behavior as a result of changes in transit service and type. Such information is essential to support further Smart Growth and TOD development strategies and to provide evidence for the need to improve the pedestrian environment around transit nodes. The distance from home or work to the nearest light rail station is known as a dose-response measure that acts as a proxy of accessibility. In addition, however, it will differentially impact individual walking behavior depending on how far people live or work from a transit station. A better understanding of the effect of distance on behavior change will help guide where development incentives for built environment change should be located near station areas.

With all these benefits to residents around light rail stations, this study examined the impact of introduction of Sound Transit light rail on walking behavior and tested the hypothesis that (1) there would be a shift in participants' walking behavior after introduction of light rail, (2) the shift is associated with the distance from home to the nearest LRT station, frequency of light rail use and drivable vehicles per adult in a household.

METHODS

Participants

Thirteen stations of Sound Transit Light Rail Service opened in fall 2009. The 14+ mile segment links Westlake Station in downtown Seattle and Sea-Tac Airport. Participants consented to participate and the study was approved by the Seattle Children's Institutional Review Board. A total of 230 participants living within 1 mile of 13 LRT stations in the present analysis were selected from the baseline(2008) and post 1 (2010) sample for the Travel Assessment and Community (TRAC) study examining the impact of a new light rail system. The distance of 1 mile from stations is the longest distance that people are willing to walk to LRT (Table 4-1) so that the study includes all potential affected participants by the introduction of LRT. Preliminary analyses compared the longitudinal sample of 214 to 62 participants who were excluded from the analysis (the 46 baseline participants who did not participate in post 1 and 16 participants who moved their residential locations). The two groups did not differ on some demographic characteristics by using a 2-sample test for equality of proportions with continuity correction. However, the longitudinal sample contained fewer people with age <40 (12% vs 35%, p<0.001), fewer people with income < 50K (35% vs 53%, p=0.016) and fewer overweight and obese people (50% vs. 70%, p<0.001).

Table 4-1 One-way distance by users to LRT

Reference	Sampling frame and process	Mean distance walked from home to LRT	Longest distance walked to LRT
Beimborn (Beimborn et al., 2003)	Portland regional travel diaries	~.24 miles	1.14 miles
Dill (Dill, 2008)	Portland residents near LRT stations	~.33 miles	~.93 miles
Kim (Kim et al., 2007)	St. Louis LRT users	.47 miles	95% walked <1.0 miles
Olszewski & Wibowo (Wibowo, S.K.Olszewski, 2005)	Interviews at Singapore LRT stations	.40 miles	Upper quartile >.5 miles
O'Sullivan & Morrall (O'Sullivan and Morrall, 1996)	Interviews at Calgary LRT stations	.40 miles	N/A
Stringham (Stringham, 1982)	Toronto residents near LRT stations	.57 miles	Upper quartile >~.67 miles
Weinstein (Weinstein Agrawal et al., 2008)	Interviews at San Francisco and Portland LRT stations	.58 miles	Upper quartile >.69 miles

Note: When only minutes of walking were reported, a 3 mins/hour(20 mins/mile) walking pace was assumed.

Data collection

Eligible and interested participants were mailed an accelerometer (Actigraph GT1M), portable GPS device (GlobalSat DG-100), and 7-day paper travel log. Participants were also provided a written or on-line (based on their preference) attitudinal and demographic survey to complete. Soon after receiving these materials, participants were contacted by study staff to review procedures (e.g., how to wear the devices; how to charge the GPS device nightly) and asked to wear the accelerometer and GPS for seven days during waking hours and to complete the travel log for these days. Accelerometer data were aggregated to 30-second epochs and GPS devices were set to collect data at 30-second intervals. Participants mailed backed the devices and travel log (and survey if written form) in a pre-paid envelope (Saelens et al., 2014).

Data processing

The process by which accelerometer data were integrated with GPS and travel log information to identify walking and non-walking physical activity bouts is described elsewhere in detail (Kang et al., 2013). Briefly, bouts of ≥5 minutes of accelerometer counts averaging >500 per 30-second epoch were considered to be physical activity; these bouts were then considered to be walking based on GPS speeds and/or on temporal overlap or proximity to walking trips recorded in the travel log. Travel logs included the places reported to be visited, the duration of each visit and the travel mode used between places. In order for an assessment day to be considered in the present analysis, it had to have at least one place record in the travel diary and an accelerometer wearing time of ≥ 8 hours per day. Accelerometer periods of ≥ 20 minutes with continuous zeros were considered as nonwearing times. An assessment day may or may not have had GPS data.

Outcome and covariates

Walking was defined as a non-mechanical and human-powered travel associated with sustained light and moderate intensity physical activity for at least 7 minutes with a 2 minutes tolerance of lower PA intensity. The definition served to isolate “walking as travel in space” from other types of PA (Kang et al., 2013).

The light rail station walking was defined as walking bouts intersecting with quarter mile buffer around 13 LRT stations (Figure 4-1). For stations that have more than one entrance, quarter mile buffer was generated for each entrance and multiple buffers for the station were merged as one station area. Bernick and Cervero (1997) list the geographical boundaries of Transit Oriented Development (TOD) as an area that extends roughly 1/4 to 1/3 of a mile from a transit station. District boundary definition in TOD ordinances in Portland and Seattle is 1/4 mile (Community Design + Architecture, 2001) and the rule of thumb is that a TOD "catchment area" spans within 1/4 mile radius of a station, or a five to seven minute walk, of a transit station (National League of Cities, 2013).

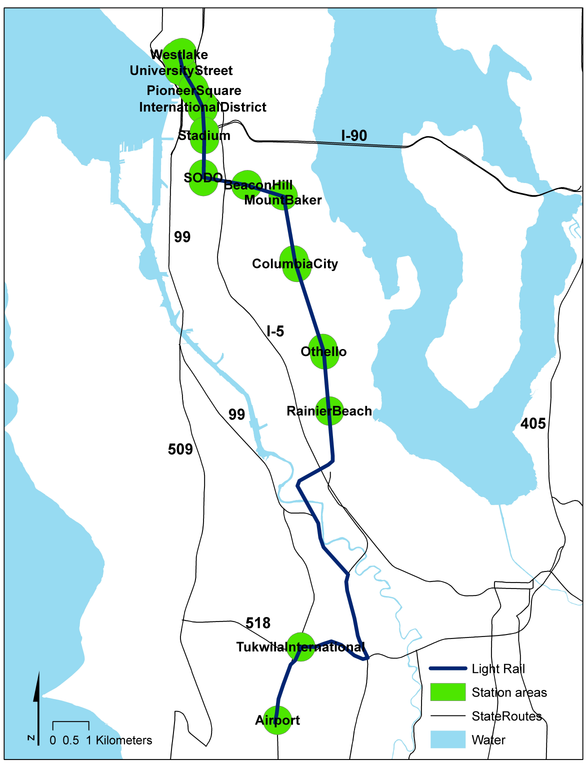

Figure 4-1 Quarter mile LRT station areas

The proportion of LRT station walking was calculate as daily LRT station walking divided by the daily overall walking (all walking bouts with GPS) of each participant, represent the spatial concentration of walking around LRT stations for each participant at baseline and post 1. At baseline walking bouts that occurred around LRT stations were non-LRT-station related, acting as a reference for post 1, while at post 1 walking bouts that occurred around LRT stations included non-LRT-station related and LRT-station related. The difference in LRT station walking between baseline and post 1 would be the effect of LRT on the residents' walking behavior. The LRT-station related walking bouts benefit from pedestrian-supportive facilities (e.g. sidewalks, curb ramps, adequate street lighting, benches, and signs) and neighborhood safety improvement by increasing pedestrians in neighborhoods near stations. After the opening of the LRT system, station areas became more supportive of pedestrians and transit users by making walking easy. LRT station areas become major activity corridors, not only being centers of specific types of land uses, but functioning as important places in the community.

Socio-demographic characteristics including age, gender, race, income and education were collected from the survey on demographics, transport and physical activity attitudes, and neighborhood perceptions. Self-reported height and weight were used to calculate BMI. Frequent light rail user was defined as 4-5 times/week from a survey question in a transportation changes section in the survey, which states as "During the average month, how often do you use the light rail?" with choices as "(1) <once a month, (2)1-2 times/month, (3) 3-4 times/month, (4) once a week, (5) 2-3 times/week (6) 4-5 times/week." Drivable motor vehicles per adult were calculated as the total number of drivable motor vehicles divided by the number of adults >18 years old in the household. These numbers were reported by participants in general information section in the survey: "How many people (including yourself) live in your household?"; "How many children under age 18 are living in your household (if any)?"; How many drivable motor vehicles (cars, trucks, motorcycles) are there available to drivers in your household?"

Distance from residential locations to the nearest LRT station was defined as shortest network route accessible by walking from home to the nearest LRT station entrance. The street entrances of LRT stations and residential locations were geocoded. Station maps of the latest SoundTransit website (http://www.soundtransit.org/Schedules/Central-Link-light-rail?tab=Stations) were used for geocoding. When entrances were not indicated on the maps, online satellite image and street view services (Google map and Bing map) were used to find entrance locations. ESRI StreetMap Premium NAVTEQ Street Data 2009 Release 1 for Washington State was used for routing and measuring network distances.

Data analysis

Descriptive statistics at baseline and post 1 on the longitudinal sample were calculated for socio-demographic characteristics, proportion with BMI in normal and overweight/obese, means and standard deviations of drivable vehicle per adult, overall walking bout minutes per person, LRT station walking bout minutes per person and proportion of LRT station walking to overall walking. The proportion of frequent light rail users was calculated for post 1.

To test whether residential location influenced LRT station walking and the proportion of LRT station walking, participants were divided into four groups based on their residential locations: 0-0.25mile, 0.25-0.5 mile, 0.5-0.75 mile and 0.75-1 mile. The average person-day level was used to calculate means and standard deviations of overall walking bout minutes, LRT station walking bouts minutes and proportion of LRT station walking bouts for baseline and post 1 and for change between the two assessment periods.

A longitudinal ordinary least squares regression was used to predict the proportion LRT station walking at post 1 using the proportion LRT station walking at baseline to control for individual difference in walking behavior at baseline. Socio-demographic characteristics age, gender, race, income, education and obesity status were also included as control variables. Distance from the residential location to the nearest light rail station, frequent light rail use and the number of drivable vehicles per adult were used as predictors for change in the proportion LRT station walking from baseline to post 1.

RESULTS

Descriptive statistics on the longitudinal panel sample are provided in Table 4-2. This sample was 63% female, 63% with age 40-65, 71% non-Hispanic White and 68% with a college degree. Income and BMI increased slightly from baseline to post 1, with about half of the sample being overweight and obese. About 7% of the participants were frequent light rail users (4-5 times per week). There was a small decrease in the number of drivable vehicles per adult from baseline to post 1. Daily minutes of overall walking bouts per person significantly decreased from baseline to post 1, while daily minutes of walking bouts near LRT stations increased. The proportion of LRT station to overall walking increased significantly from baseline to post 1.

Table 4-2 Descriptive statistics for panel sample (n=214)

		Baseline		Post 1		
		Mean	SD	Mean	SD	p-value
Gender	male	0.37		0.37		
	female	0.63		0.63		
Age	18-39	0.16		0.12		
	40-65	0.66		0.65		
	>65	0.18		0.24		
Race	Hispanic or non-White	0.29		0.29		
	Non-Hispanic White	0.71		0.71		
Income	<50k	0.37		0.36		
	50-100k	0.40		0.38		
	>100k	0.23		0.26		
Education	Less than college graduate	0.30		0.30		
	college graduate	0.70		0.70		
BMI	BMI < 25	0.51		0.49		
	BMI >= 25	0.49		0.51		
LRT user	non-frequent LRT user	-		0.93		
	Frequent LRT user	-		0.07		
Drivable vehicle	vehicle/adult	0.75	0.54	0.74	0.53	0.860
Walking	Overall walking (minutes/person-day)	28.04	26.67	23.67	22.71	0.002
	LRT station walking (1/4 mile) (minutes/person-day)	6.56	11.25	7.83	13.45	0.113
	Proportion LRT station walking (%)	0.19	0.27	0.25	0.30	0.014

A sensitivity test grouping participants living within different rings around LRT stations showed that total minutes of daily walking decreased for all groups except that living closest to the nearest station entrance (¼ mi) (Table 4-3, Figure 4-2). The mean change in minutes of daily LRT station walking from baseline to post 1 decreased with distance from LRT stations, with participants living between 3/4 and one mile walking less in post 1 than at baseline. The mean change of percent of LRT station walking from baseline to post 1 was largest in rings 0.25-0.50 and 0.5-0.75.

Table 4-3 Sensitivity of distance from home to LRT stations (n=214)

	Overall walking (minutes/person-day)		**LRT station walking (1/4 mile)**		**% LRT station walking (average over individual %s)**	
	Mean	SD	Mean	SD	Mean	SD
	Distance from home to the nearest LRT station: 0-0.25mile (n=26)					
baseline	27.54	27.31	9.37	11.94	35.7%	39.1%
post1	27.83	28.60	14.70	20.02	41.6%	35.5%
Change	0.29	20.86	5.33	19.74	6.0%	50.3%
	Distance from home to the nearest LRT station: 0.25-0.50mile (n=61)					
baseline	32.40	28.67	9.26	13.59	25.8%	28.6%
post1	27.19	21.15	10.87	15.18	33.1%	32.2%
Change	-5.21	20.59	1.61	13.78	7.2%	28.6%
	Distance from home to the nearest LRT station: 0.50-0.75mile (n=73)					
baseline	22.71	24.18	3.30	6.23	10.8%	19.0%
post1	17.56	16.99	4.19	7.17	18.9%	28.8%
Change	-5.16	18.26	0.89	8.37	8.2%	32.9%
	Distance from home to the nearest LRT station: 0.75-1.00mile (n=70)					
baseline	29.76	26.69	6.49	12.00	16.3%	23.4%
post1	25.29	25.92	6.39	12.86	16.9%	24.3%
Change	-4.46	23.80	-0.09	8.09	0.6%	24.1%

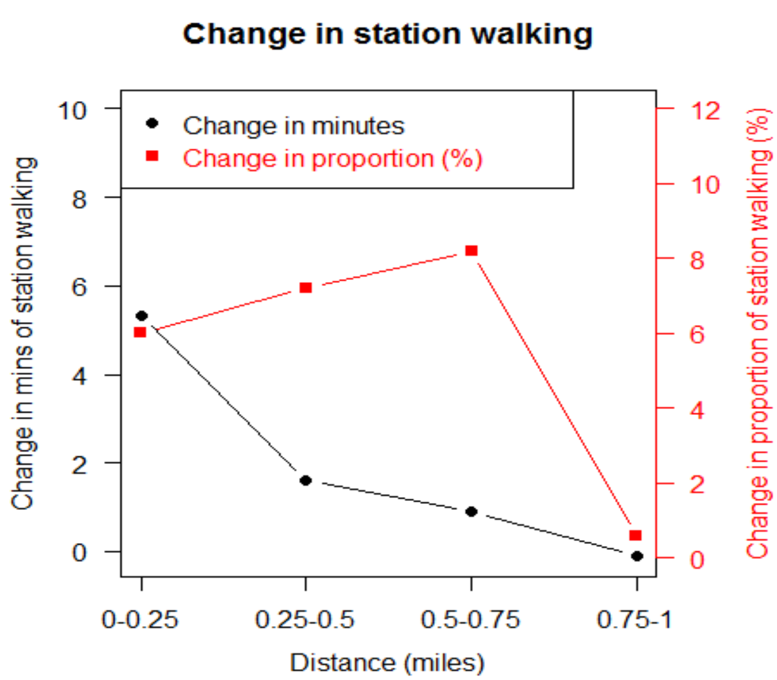

Figure 4-2 Change in station walking vs. Distance from homes to stations

The longitudinal analysis (Table 4-4) shows that the proportion of LRT station walking at baseline was positively related to that of post 1. After adjusting for the proportion of LRT station walking at baseline, college graduates and frequent light rail users significantly increased their proportion of LRT station walking in post 1; overweight and obese participants significantly decreased their proportion of LRT walking. The distance from the residential location to the nearest light rail station was negatively related to the change in the proportion of LRT station daily walking minutes.

Table 4-4 Longitudinal OLS to predict proportion of LRT station walking at post 1

	Estimate	Std. Error	p-value	
(Intercept)	0.274	0.092	0.003	**
Proportion LRT station walking (baseline)	0.334	0.078	<0.001	***
Male	0.081	0.044	0.065	
Age 40-65	0.026	0.067	0.696	
Age >65	-0.005	0.074	0.944	
Non-Hispanic White	-0.025	0.049	0.616	
College graduate	0.157	0.047	0.001	**
Income 50-100K	0.020	0.051	0.700	
Income >100K	-0.072	0.058	0.215	
Overweight and obese	-0.090	0.042	0.033	*
Distance to nearest station	-0.264	0.089	0.004	**
Use LRT 4-5 times/week	0.237	0.080	0.003	**
Drivable vehicle per adult	-0.034	0.042	0.413	

*P<0.05, **P<0.01, ***P<0.001

DISCUSSION

Overall, daily minutes of walking decreased from baseline to post 1 while walking around stations increased, which lead to a significant increase in the proportion of walking that took place around LRT stations. The decrease in overall walking might be due to aging, since there is a 5% decrease in the number of participants aged 18-29 and a 5% increase in the above 65 age group. On the other hand, the increase in the proportion of walking minutes around LRT station areas suggests a shift in the geographic location of walking from outside to inside the station areas after the introduction of light rail for participants who have convenient (<one mile) access to light rail. This walking behavior shift in location supports previous evidence that improved accessibility to transit and specifically rail transit promotes walking near transit stations. Rail transit thus helps increase the number of potential customers for retail and other service activities near station areas (Besser and Dannenberg, 2005; MacDonald et al., 2010).

Distance between home and the nearest light rail station plays an important role in the shift of geographic location for walking. The negative relation between the distance to the nearest station and the change of walking minutes around LRT stations appears to be strongest for distances <0.75 mile. This implies that target TOD development areas should focus on station service areas within <0.75 mile walking distance.

Beyond distance from home to the station area, the shift of individual walking into LRT station areas is influenced by weight status, and light rail usage frequency, after adjusting for socio-demographic characteristics. Compared to respondents with normal weight, the overweight and obese tended to decrease the proportion of walking around station areas, implying that the transportation infrastructure changes had less impact on populations who might have developed sedentary lifestyle and became less sensitive to facilities supporting active living.

After controlling for socio-demographic characteristics, the increased proportion of walking around station areas is significantly negatively related to distance from home to the nearest station, which is not only due to the ease of access to light rail service but also to pedestrian features and possible change of neighborhood perception for people living close to LRT stations. The increased proportion of walking around LRT stations among frequent light rail users indicates that walking to, from or around stations might substitute some walking outside of ¼ mile station areas. Also frequent rail users may have better knowledge of various destinations and environmental changes along their walking paths (and hence want to use these destinations more frequently) than those who use light rail less since introduction of light rail.

The strengths of the present study include the objective measure of walking and the natural experiment study design. Since walking was objectively measured, it becomes feasible to study location-based walking behavior, i.e. walking around LRT stations. The effect of self-selection (i.e., residents selecting neighborhoods having environments consistent with their activity preferences), considered as a key issue in determining the direction of causal pathways between the built environment and travel behavior, is minimized in a prospective cohort design examining the same individuals' walking behavior before and after introduction of light rail (i.e., a natural experiment). Furthermore, the study provides a rare pre- and post-test of a new LRT line intervention with large sample size, which suggests new planning, research and policy directions potentially influencing transportation planning decisions.

Like many studies using wearable devices in quasi-experimental intervention design, the study has limitations. First, the 7-day assessment period is short and may not represent free-living daily activity and some of the walking across the ¼ mi area of an LRT station may have occurred by chance. Second, the sample was drawn from only one U.S. metropolitan region, restricting the generalizability of the findings. This study is part of a large project investigating the impact of light rail on travel behavior and physical activity, and participants had more transit access than the general U.S. population. Finally, physical activity attitudes and neighborhood perceptions may have effect on the walking behavior around home neighborhood and need to be considered in the future studies.

CONCLUSION

TOD starts with the introduction of station in a neighborhood. This longitudinal study provides evidence to illustrate the causal relation between LRT and change in walking behavior and to emphasize the important of distance in the modes of travel used in TOD. The walking behavior change represented by the increase in walking around station areas validates the multiple goals of TOD (providing housing near transit stations, destinations attractive to pedestrian, and integrating physical activity into everyday activity) and supports the TOD strategy of building new home and business proximate to light rail stations.TOD can help achieve health and environmental goals.

Chapter 5 : CONCLUSION

The comprehensive spatial analysis in chapter 2 disentangled the complex role that the local environment plays in obesity, and more generally in health-supportive environments. The analyses showed that people cluster in space by BMI status. The two different cluster detection techniques used in the study yielded similar results on spatial clustering by BMI while relying on different assumptions and mechanisms to identify spatial clustering. These results implied that spatial clustering by BMI is not biased by the methods used to detect clustering. Furthermore, the individual-level analyses represented a departure from research using aggregated geographies and helped consider the fine-grained characteristics of environments where participants reside. Also, the use of residuals from regression models as the outcome of the spatial analyses was essential to detect the spatial dependence structure for inference (Pullan et al., 2012). The findings suggested that two characteristics of the home neighborhood environment, neighborhood property values and residential density, might serve to identify the locations of obesity prevention and intervention programs.

The cross-sectional study in chapter 3 fills a gap in the literature by using objective data on the frequency and location of walking assessed over the course of 7 days. These data are first to preview habitual walking as it occurs inside or outside of the home neighborhood. The study presents evidence that most walking occurs in the home neighborhood, thus confirming the value of the many past studies focusing on the influence of the home neighborhood environment on walking activity. In addition, walking near home appears to be significantly more frequent in neighborhoods of higher residential and employment density. However, results show that only those neighborhoods with the highest levels of residential density or employment density would support higher amounts of walking.

The study in chapter 4 also used objective measure of walking, but expanded the examination of walking in non-residential settings and considered changes in walking longitudinally. Measuring walking objectively and tracking it with GPS makes it feasible to examine location-based walking behavior, and in this case, walking around LRT stations. In addition, considering changes in walking near LRT stations minimized the effect of self-selection (i.e., residents selecting neighborhoods having environments consistent with their activity preferences), considered as a key issue in determining the direction of causal pathways between the built environment and travel behavior. Thus the use of a prospective cohort design examining the same individuals' walking behavior before and after the introduction of light rail (i.e., a natural experiment) helps determining the potential causal effects of environment on walking behavior. The study found that the overall walking decreased from baseline to post 1, but increasing walking occurred around stations, which indicated there was a substitution of walking around stations and suggested LRT increased the number of potential customers for retail and other service activities near station areas. Finally, the study provides a rare pre- and post-test of a new LRT line intervention with large sample size, which suggests new planning, research and policy directions potentially influencing transportation planning decisions.

Future studies need to investigate the effect of distance thresholds selected for the spatial weights in the Local Moran's I and the maximum cluster size in spatial scan statistic on the results of cluster detection. In the study in Chapter 2, the 2 miles distance was hypothesis-driven and would needs to be tested in similar obesity studies. In addition, methodological comparison should focus on what is the impact of sampling scheme, e.g. random sample or stratified sample, on the performance of spatial detection techniques. Finally, since the circular buffers used to calculate the neighborhood variables might not represent actual neighborhood boundaries, sophisticated spatial detection methods need to be invented to take into account complex shapes for neighborhoods.

The ability to demonstrate the relation between built environment and walking behavior relies on accurate measures in time and space. More future studies need to combine advance technologies such as GPS, accelerometer and GIS to reveal the dynamic interaction between the built environment and travel behaviors. Further analysis needs to be done on data loss including missing GPS signal, cold start and non-wearing device, because all these problems may introduce biases into the measurement and thus influence the study results. Reducing participant burden by using advanced technologies in smart phone would not only resolve the data loss problem to some extent, but also make longer assessment study feasible and thus minimize the number of dropouts. Finally, researchers should expand the studies into non-residential settings to understand the relationship of a wide array of built environment characteristics and would provide critical insights into the effects on travel behaviors.

BIBLIOGRAPHY

Anselin, L., 1995. Local Indicators of Spatial Association. Geogr. Anal., 2 27, 93–115.

Anselin, L., Syabri, I., Kho, Y., 2006. GeoDa: An Introduction to Spatial Data Analysis. Geogr. Anal. 38, 5–22. doi:10.1111/j.0016-7363.2005.00671.x

Beimborn, E., Greenwald, M., Jin, X., 2003. Accessibility, Connectivity, and Captivity: Impacts on Transit Choice. Transp. Res. Rec. doi:10.3141/1835-01

Belsley, D.A., 1991. Conditioning diagnostics : collinearity and weak data in regression. Wiley, New York.

Besser, L.M., Dannenberg, A.L., 2005. Walking to public transit: steps to help meet physical activity recommendations. Am. J. Prev. Med. 29, 273–280.

Block, J.P., Christakis, N.A., O'Malley, A.J., Subramanian, S. V, 2011. Proximity to food establishments and body mass index in the Framingham Heart Study offspring cohort over 30 years. Am. J. Epidemiol. 174, 1108–14. doi:10.1093/aje/kwr244

Brown, B.B., Werner, C.M., 2007. A new rail stop: tracking moderate physical activity bouts and ridership. Am. J. Prev. Med. 33, 306–309.

Brown, B.B., Werner, C.M., 2008. Before and After a New Light Rail Stop: Resident Attitudes, Travel Behavior, and Obesity. J. Am. Plan. Assoc. doi:10.1080/01944360802458013

Brown, B.B., Werner, C.M., 2011. The Residents' Benefits and Concerns Before and After a New Rail Stop: Do Residents Get What They Expect? Environ. Behav. doi:10.1177/0013916510392030

Brownson, R.C., Boehmer, T.K., Luke, D.A., 2005. DECLINING RATES OF PHYSICAL ACTIVITY IN THE UNITED STATES: What Are the Contributors? Annu. Rev. Public Health 26.

Calthorpe, P., 1993. The next American metropolis : ecology, community, and the American dream. Princeton Architectural Press, New York.

Cameron Trivedi, P. K., A.C., 1998. Regression analysis of count data. Cambridge University Press, Cambridge, UK; New York, NY, USA.

Cameron, A.J., Crawford, D.A., Salmon, J., Campbell, K., McNaughton, S.A., Mishra, G.D., Ball, K., 2011. Clustering of obesity-related risk behaviors in children and their mothers. Ann. Epidemiol. 21, 95–102. doi:10.1016/j.annepidem.2010.11.001

Centers for Disease Control and Prevention, n.d. Obesity prevalence in 2012 varies across states and regions [WWW Document]. URL http://www.cdc.gov/obesity/data/adult.html#Prevalence (accessed 1.22.14).

Cervero, R., 1994. Transit-based housing in California: evidence on ridership impacts. Transp. Policy 1, 174–183. doi:10.1016/0967-070X(94)90013-2

Cervero, R., 2001. Walk-and-Ride: Factors Influencing Pedestrian Access to Transit. J. Public Transp. 3, 1–23.

Cervero, R., 2006. Alternative Approaches to Modeling the Travel-Demand Impacts of Smart Growth. J. Am. Plan. Assoc. 72, 285–295. doi:10.1080/01944360608976751

Chang, V.W., Christakis, N.A., 2005. Income inequality and weight status in US metropolitan areas. Soc. Sci. Med. 61, 83–96. doi:10.1016/j.socscimed.2004.11.036

Christakis, N.A., Fowler, J.H., 2007. The spread of obesity in a large social network over 32 years. N. Engl. J. Med. 357, 370–379. doi:10.1056/NEJMsa066082

Coogan, P.F., White, L.F., Adler, T.J., Hathaway, K.M., Palmer, J.R., Rosenberg, L., 2009. Prospective study of urban form and physical activity in the Black Women's Health Study. Am J Epidemiol 170, 1105–1117.

Crépey, P., Barthélemy, M., 2007. Detecting Robust Patterns in the Spread of Epidemics: A Case Study of Influenza in the United States and France. Am. J. Epidemiol. 166, 8.

Diez-Roux, A. V, 1998. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. Am. J. Public Health 88, 216–222.

Dill, J., 2008. Transit Use at Transit-Oriented Developments in Portland, Oregon, Area. Transp. Res. Rec. J. Transp. Res. Board. doi:10.3141/2063-19

Drewnowski, A., Aggarwal, A., Hurvitz, P.M., Monsivais, P., Moudon, A.V., 2012. Obesity and Supermarket Access: Proximity or Price? Am. J. Public Health 102, e1–e7. doi:10.2105/AJPH.2012.300660

Drewnowski, A., Moudon, A.V., Jiao, J., Aggarwal, A., Charreire, H., Chaix, B., 2013. Food environment and socioeconomic status influence obesity rates in Seattle and in Paris. Int. J. Obes. 1–9. doi:10.1038/ijo.2013.97

Drewnowski, A., Rehm, C.D., Solet, D., 2007. Disparities in obesity rates: analysis by ZIP code area. Soc. Sci. Med. 65, 2458–2463.

Drewnowski, A., Specter, S., 2004. Poverty and obesity: the role of energy density and energy costs. Am. J. Clin. Nutr. 79, 6–16.

Duncan, M.J., Winkler, E., Sugiyama, T., Cerin, E., duToit, L., Leslie, E., Owen, N., 2010. Relationships of land use mix with walking for transport: do land uses and geographical scale matter? J Urban Heal. 87, 782–795. doi:10.1007/s11524-010-9488-7

Ewing, R., Schmid, T., Killingsworth, R., Zlot, A., Raudenbush, S., 2008. Relationship Between Urban Sprawl and Physical Activity, Obesity, and Morbidity. Urban Ecol. 18, 567–582.

Eyler, A.A., Brownson, R.C., Bacak, S.J., Housemann, R.A., 2003. The epidemiology of walking for physical activity in the United States. Med. Sci. Sports Exerc. 35, 1529–1536.

Feng, J., Glass, T.A., Curriero, F.C., Stewart, W.F., Schwartz, B.S., 2010. The built environment and obesity: a systematic review of the epidemiologic evidence. Health Place 16, 175–190.

Forsyth, A., Oakes, J.M., Schmitz, K.H., Hearst, M., 2007. Does Residential Density Increase Walking and Other Physical Activity? Urban Stud. 44, 679–697. doi:10.1080/00420980601184729

Frank, L., Kavage, S., Litman, T., 2006. Promoting public health through Smart Growth: Building healthier communities through transportation and land use policies and practices. Smart Growth BC 1–43.

Frank, L.D., Pivo, G., 1994. Impacts of mixed use and density on utilization of three modes of travel: Single occupant vehi- cle, transit, and walking. Transp. Res. Rec. J. Transp. Res. Board 1466, 44–52.

Frank, L.D., Saelens, B.E., Powell, K.E., Chapman, J.E., 2007. Stepping towards causation: do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? Soc. Sci. Med. 65, 1898–914. doi:10.1016/j.socscimed.2007.05.053

Frank, L.D., Schmid, T.L., Sallis, J.F., Chapman, J.E., Saelens, B.E., 2005. Linking objectively measured physical activity with objectively measured urban form - Findings from SMARTRAQ. Am. J. Prev. Med. 28, 117–125.

Gallos, L.K., Barttfeld, P., Havlin, S., Sigman, M., Makse, H.A., 2012. Collective behavior in the spatial spreading of obesity. Sci. Rep. doi:10.1038/srep00454

Gesink, D.C., Sullivan, A.B., Miller, W.C., Bernstein, K.T., 2011. Sexually transmitted disease core theory: roles of person, place, and time. Am. J. Epidemiol. 174, 81–89.

Getis, A., Ord, J.K., 1996. Local spatial statistics: an overview, in: Longley, P., Batty, M. (Eds.), Spatial Analysis Modelling in a GIS Environment. GeoInformation International, Distributed in the Americas by J. Wiley, pp. 261–277.

Gordon-Larsen, P., Nelson, M.C., Page, P., Popkin, B.M., 2006. Inequality in the built environment underlies key health disparities in physical activity and obesity. Pediatrics 117, 417–424. doi:10.1542/peds.2005-0058

Greene, S.K., Ionides, E.L., Wilson, M.L., 2006. Patterns of influenza-associated mortality among US elderly by geographic region and virus subtype, 1968-1998. Am. J. Epidemiol. 163, 316–326.

Hill Peters, J C, J.O., 1998. Environmental Contributions to the Obesity Epidemic. Science. 280, 1371.

Huang, L., Stinchcomb, D.G., Pickle, L.W., Dill, J., Berrigan, D., 2009. Identifying clusters of active transportation using spatial scan statistics. Am. J. Prev. Med. 37, 157–66. doi:10.1016/j.amepre.2009.04.021

Hurvitz, P.M., Moudon, A.V., 2012. Home Versus Nonhome Neighborhood: Quantifying Differences in Exposure to the Built Environment. Am. J. Prev. Med. 42.

Hurvitz, P.M., Moudon, A.V., Kang, B., Saelens, B.E., Duncan, G.E., 2014. Emerging Technologies for Assessing Physical Activity Behaviors in Space and Time. Front. Public Heal. 2. doi:10.3389/fpubh.2014.00002

Jacquez, G.M., Greiling, D.A., 2003. Local clustering in breast, lung and colorectal cancer in Long Island, New York. Int. J. Health Geogr. 12, 1–12.

Jiao, J., Moudon, A.V., Drewnowski, A., 2010. Grocery shopping: How individuals and built environments influence travel mode choice. Transp. Res. Rec.

Kang, B., Moudon, A.V., Hurvitz, P.M., Reichley, L., Saelens, B.E., 2013. Walking Objectively Measured: Classifying Accelerometer Data with GPS and Travel Diaries. Med. Sci. Sports Exerc. 45, 1419–28. doi:10.1249/MSS.0b013e318285f202

Kawachi, I., Berkman, L.F., 2003. Neighborhoods and health. Oxford University Press, Oxford; New York.

Kelly-Schwartz, A.C., Stockard, J., Doyle, S., Schlossberg, M., 2004. Is sprawl unhealthy? A multilevel analysis of the relationship of metropolitan sprawl to the health of individuals. J. Plan. Educ. Res. 24, 184–196.

Kim, S., Ulfarsson, G.F., Todd Hennessy, J., 2007. Analysis of light rail rider travel behavior: Impacts of individual, built environment, and crime characteristics on transit access. Transp. Res. Part A 41, 511–522.

Kitron, U., Kazmierczak, J.J., 1997. Spatial analysis of the distribution of Lyme disease in Wisconsin. Am. J. Epidemiol. 145, 558–566.

Kligerman, M., Sallis, J.F., Ryan, S., Frank, L.D., Nader, P.R., 2007. Association of neighborhood design and recreation environment variables with physical activity and body mass index in adolescents. Am. J. Heal. Promot. 21, 274–277.

Kulldorff, M., 1997. A spatial scan statistic. Commun. Stat. Theory Methods 26, 1481–1496. doi:10.1080/03610929708831995

Kulldorff, M., Feuer, E.J., Miller, B.A., Freedman, L.S., 1997. Breast cancer clusters in the northeast United States: a geographic analysis. Am. J. Epidemiol. 146, 161–170.

Kulldorff, M., Huang, L., Pickle, L., Duczmal, L., 2006. An elliptic spatial scan statistic. Stat. Med. 25, 3929–3943.

Lee, C., Moudon, A.V., 2006a. Correlates of Walking for Transportation or Recreation Purposes. J. Phys. Act. Heal. 3, S77–S98.

Lee, C., Moudon, A.V., 2006b. The 3Ds + R: Quantifying land use and urban form correlates of walking. Transp. Res. Part D Transp. Environ. 11, 204–215.

Lee, P.H., Macfarlane, D.J., Lam, T., Stewart, S.M., 2011. Validity of the international physical activity questionnaire short form (IPAQ-SF): A systematic review. Int. J. Behav. Nutr. Phys. Act. doi:10.1186/1479-5868-8-115

Li, F., Harmer, P., Cardinal, B.J., Bosworth, M., Johnson-Shelton, D., 2009. Obesity and the built environment: does the density of neighborhood fast-food outlets matter? The science of health promotion. 23, 203–209.

Ludwig, J., Sanbonmatsu, L., Gennetian, L., Adam, E., Duncan, G., Katz, L., Kessler, R., Kling, J., Lindau, S., Whitaker, R., McDade, T., 2011. Neighborhoods, obesity, and diabetes - A randomized social experiment. N. Engl. J. Med. 365, 1509–1519.

MacDonald, J.M., Stokes, R.J., Cohen, D.A., Kofner, A., Ridgeway, G.K., 2010. The Effect of Light Rail Transit on Body Mass Index and Physical Activity. Am. J. Prev. Med. 39, 105–112.

McConville, M.E., Rodriguez, D.A., Clifton, K., Cho, G., Fleischhacker, S., 2011. Disaggregate land uses and walking. Am. J. Prev. Med. 40, 25–32. doi:10.1016/j.amepre.2010.09.023

Michimi, A., Wimberly, M.C., 2010. Spatial patterns of obesity and associated risk factors in the conterminous U.S. Am. J. Prev. Med. 39, e1–e12. doi:10.1016/j.amepre.2010.04.008

Mokdad, A.H., Ford, E.S., Bowman, B.A., Dietz, W.H., Vinicor, F., Bales, V.S., Marks, J.S., 2003. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. JAMA J. Am. Med. Assoc. 289, 76–79. doi:10.1001/jama.289.1.76

Moore, K., Diez Roux, A. V, Auchincloss, A., Evenson, K.R., Kaufman, J., Mujahid, M., Williams, K., 2013. Home and work neighbourhood environments in relation to body mass index: the Multi-Ethnic Study of Atherosclerosis (MESA). J. Epidemiol. Community Health 67, 846–53. doi:10.1136/jech-2013-202682

Morris, J.N., Hardman, A.E., 1997. Walking to health. Sports Med. 23, 306–332. doi:10.2165/00007256-199723050-00004

Moudon, A.V., Cook, A.J., Ulmer, J., Hurvitz, P.M., Drewnowski, A., 2011a. A neighborhood wealth metric for use in health studies. Am. J. Prev. Med. 41, 88–97.

Moudon, A.V., Lee, C., Cheadle, A.D., Garvin, C., Rd, D.B.J., Schmid, T.L., Weathers, R.D., 2007. Attributes of environments supporting walking. Am. J. Heal. Promot.

Moudon, A.V., Sohn, D.W., Kavage, S., Mabry, J., 2011b. Transportation-Efficient Land Use Mapping Index (TELUMI), a Tool to Assess Multimodal Transportation Options in Metropolitan Regions. Int. J. Sustain. Transp. 5, 111–133. doi:10.1080/15568311003624262

National League of Cities, 2013. Transit- Oriented Development (TOD) [WWW Document]. URL http://www.sustainablecitiesinstitute.org/topics/land-use-and-planning/transit-oriented-development-(tod) (accessed 6.26.14).

Norman, G.J., Carlson, J.A., O'Mara, S., Sallis, J.F., Patrick, K., Frank, L.D., Godbole, S. V., 2013. Neighborhood Preference, Walkability and Walking in Overweight/Obese Men. Am. J. Health Behav. doi:10.5993/AJHB.37.2.15

O'Sullivan, S., Morrall, J., 1996. Walking Distances to and from Light-Rail Transit Stations. Transp. Res. Rec. doi:10.3141/1538-03

Ogilvie, D., Foster, C.E., Rothnie, H., Cavill, N., Hamilton, V., Fitzsimons, C.F., Mutrie, N., 2007. Interventions to promote walking: systematic review. BMJ 334, 1204.

Omer, S.B., Enger, K.S., Moulton, L.H., Halsey, N.A., Stokley, S., Salmon, D.A., 2008. Geographic clustering of nonmedical exemptions to school immunization requirements and associations with geographic clustering of pertussis. Am. J. Epidemiol. 168, 1389–1396.

Papas, M.A., Alberg, A.J., Ewing, R., Helzlsouer, K.J., Gary, T.L., Klassen, A.C., 2007. The built environment and obesity. Epidemiol. Rev. 29, 129–143.

Patil, G.P., Taillie, C., 2004. Upper level set scan statistic for detecting arbitrarily shaped hotspots. Environ. Ecol. Stat. 11, 183–197. doi:10.1023/B:EEST.0000027208.48919.7e

Penney, T.L., Rainham, D.G.C., Dummer, T.J.B., Kirk, S.F.L., 2013. A spatial analysis of community level overweight and obesity. J. Hum. Nutr. Detetics Off. J. Br. Diet. Assoc. 1–10. doi:10.1111/jhn.12055

Pullan, R.L., Sturrock, H.J.W., Soares Magalhães, R.J., Clements, A.C.A., Brooker, S.J., 2012. Spatial parasite ecology and epidemiology: a review of methods and applications. Parasitology 139, 1870–87.

Rappaport, J., Seidman, E. (Eds.), 2000. Handbook of community psychology. New York : Kluwer Academic/Plenum, New York.

Rehm, C.D., Moudon, A.V., Hurvitz, P.M., Drewnowski, A., 2012. Residential property values are associated with obesity among women in King County, WA, USA. Soc. Sci. Med.

Rundle, A., Neckerman, K.M., Freeman, L., Lovasi, G.S., Purciel, M., Quinn, J., Richards, C., Sircar, N., Weiss, C., 2009. Neighborhood Food Environment and Walkability Predict Obesity in New York City. Environ. Health Perspect. 117, 442–447.

Rzewnicki, R., Vanden Auweele, Y., De Bourdeaudhuij, I., 2003. Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. Public Health Nutr. 6, 299–305. doi:10.1079/PHN2002427

Sabel, C.E., Boyle, P.J., Löytönen, M., Gatrell, A.C., Jokelainen, M., Flowerdew, R., Maasilta, P., 2003. Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death. Am. J. Epidemiol. 157, 898–905.

Saelens, B.E., Handy, S.L., 2008. Built environment correlates of walking: a review. Med. Sci. Sports Exerc. 40, S550–566. doi:10.1249/MSS.0b013e31817c67a4

Saelens, B.E., Moudon, A.V., Kang, B., Hurvitz, P.M., Zhou, C., 2014. Relation between higher physical activity and public transit use. Am. J. Public Health 104, 854–859.

Schuurman, N., Peters, P.A., Oliver, L.N., 2009. Are obesity and physical activity clustered? A spatial analysis linked to residential density. Obesity (Silver Spring). 17, 2202–9. doi:10.1038/oby.2009.119

Shih, M., Dumke, K.A., Goran, M.I., Simon, P.A., 2013. The association between community-level economic hardship and childhood obesity prevalence in Los Angeles. Pediatr. Obes. 8, 411–7. doi:10.1111/j.2047-6310.2012.00123.x

Siegel, P.Z., Brackbill, R.M., Heath, G.W., 1995. The epidemiology of walking for exercise: implications for promoting activity among sedentary groups. Am. J. Public Health 85, 706–710.

Stopher, P.R., Greaves, S.P., 2007. Household travel surveys: Where are we going? Transp. Res. Part A Policy Pract. doi:10.1016/j.tra.2006.09.005

Stringham, M., 1982. Travel Behavior Associated with Land Uses Adjacent to Rapid Transit Stations. ITE J. 52, 3.

Takahashi, K., Kulldorff, M., Tango, T., Yih, K., 2008. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. Int. J. Health Geogr. 7.

Takahashi, K., Kulldorff, M., Tango, T., Yih, K., 2008. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. Int. J. Health Geogr. 7, 14.

Vuong, Q.H., 1989. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. Econom. J. Econom. Soc. 57, 307–333. doi:10.2307/1912557

Waller, L.A., Gotway, C.A., 2004. Applied Spatial Statistics for Public Health Data, Environmental Health, Applied Spatial Statistics for Public Health Data. Wiley-Interscience. doi:10.1198/jasa.2005.s15

Waller, L.A., Gotway, C.A., 2004. Applied spatial statistics for public health data. John Wiley & Sons, Hoboken, N.J.

Wang, Y., Beydoun, M.A., 2007. The obesity epidemic in the United States--gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. Epidemiol. Rev. 29, 6–28. doi:10.1093/epirev/mxm007

Washington State Behavioral Risk Factor Surveillance System [BRFSS]. Questionnaire, forms A and B [WWW Document], 2008. URL http://www.doh.wa.gov/Portals/1/Documents/5300/BRFSSQuestion08A.pdf, http://www.doh.wa.gov/Portals/1/Documents/5300/BRFSSQuestion08B.pdf

Weinstein Agrawal, A., Schlossberg, M., Irvin, K., 2008. How Far, by Which Route and Why? A Spatial Analysis of Pedestrian Preference. J. Urban Des. doi:10.1080/13574800701804074

Wibowo, S.K.Olszewski, P., 2005. Modeling Walking Accessibility to Public Transport Terminals: Case Study of Singapore Mass Rapid Transit. J. East. Asia Soc. Transp. Stud. 6, 147–156.

World Health Organization (WHO), 2007. Increasing Physical Activity Reduces Risk of Heart Disease and Diabetes. Geneva:WHO.

World Health Organization (WHO), 2009. Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks, Nonserial Publication Series.

World Health Organization (WHO), 2010. Global recommendations on physical activity for health.

World Health Organization (WHO), 2012. Monitoring framework and targets for the prevention and control of NCDs [WWW Document]. URL http://www.who.int/nmh/events/2012/consultation_april_2012/en/index.html (accessed 2.5.14).