Published in final edited form as:

Epidemiology. 2017 May ; 28(3): 403–411. doi:10.1097/EDE.0000000000000631.

Methods to study variation in the associations between food store availability and body mass in the Multi-Ethnic Study of Atherosclerosis

Jonggyu Baek^a, Jana A. Hirsch^b, Kari Moore^c, Loni Philip Tabb^c, Tonatiuh Barrientos-Gutierrez^d, Lynda D. Lisabeth^a, Ana V Diez-Roux^c, and Brisa N. Sánchez^a

^aUniversity of Michigan, Ann Arbor, Michigan, USA

^bUniversity of North Carolina at Chapel Hill, North Carolina, USA

^cDrexel University, Philadelphia, Pennsylvania, USA

^dNational Institute of Public Health, Cuernavaca, Morales, Mexico

Abstract

Research linking characteristics of the neighborhood environment to health has relied on traditional regression methods where pre-specified distances from participant's locations or areas are used to operationalize neighborhood-level measures. Since the relevant spatial scale of neighborhood environment measures may differ across places or individuals, using pre-specified distances could result in biased association estimates or efficiency losses. We use novel hierarchical distributed lag models and data from the Multi-ethnic Study of Atherosclerosis (MESA) to: 1) examine whether and how the association between the availability of favorable food stores and body mass index (BMI) depends on continuous distance from participant locations (instead of traditional buffers), thus allowing us to indirectly infer the spatial scale at which this association operates; 2) examine if the spatial scale and magnitude of the association differs across six MESA sites and 3) across individuals. As expected, we found that the association between higher availability of favorable food stores within closer distances from participant's residential location was stronger than at farther distances, and that the magnitude of the adjusted association declined quickly from zero to two miles. Furthermore, between-individual heterogeneity in the scale and magnitude of the association was present; the extent of this heterogeneity was different across the MESA sites. Individual heterogeneity was partially explained by sex. This study illustrated novel methods to examine how neighborhood environmental factors may be differentially associated with health at different scales, providing nuance to previous research that ignored the heterogeneity found across individuals and contexts.

Keywords

built environment; food environment; obesity; hierarchical distributed lag model

Correspondence: Brisa N. Sánchez, PhD, Associate Professor of Biostatistics, Biostatistics Department, SPH II Room 4164, The University of Michigan School of Public Health, brisa@umich.edu, Tel. (734) 763-2451, Fax. (734) 763-2215.

Conflict of interest: The authors report no conflict of interest.

Neighborhoods have received attention as contributors to the obesity epidemic^{1,2} because neighborhood resources may constrain individuals' health behaviors³⁻⁵. Specific physical features of the neighborhood environment, such as availability of specific types of food stores⁶⁻¹¹, have been examined as potential contributors to obesity. Despite increased evidence of the relationship between the neighborhood environment and obesity, several methodological challenges remain¹². One methodological obstacle is the selection of spatial scale used to operationalize neighborhood measures^{13,14}.

Differences across studies in the spatial scale or unit used to define neighborhood environment metrics make it challenging to compare and synthesize results¹³. There is no standard geographic shape or size within which measures of the built environment attributes are constructed. Some studies use a circular buffer around participants' residence while others use census tract, block groups, counties. However, since the mechanistically relevant spatial scale is unknown, it is usually chosen in an ad hoc manner (e.g., selecting a buffer size 1/2- or 1-mile radius). Incorrect selection of the spatial scale can impact the estimated associations, including severe bias in the association of interest^{15,16}. Shifts in results when altering the size or shape of the spatial unit used to examine associations is a well-recognized problem^{17,18}. Indeed, previous work has suggested that the association between the neighborhood environment and obesity is different depending on the chosen spatial scale¹⁹⁻²¹.

An additional complication is the extent to which the relevant spatial scale of contextual factors depends on cities, neighborhoods or individuals. For instance, broad differences in built environment characteristics between cities (e.g., transportation systems) may confer differential accesses to food stores, potentially resulting in heterogeneity of the food environment-health associations across cities. Additionally, characteristics of individuals' residential locations or neighborhood or the individuals' characteristics may also alter the spatial scale. For example, street connectivity near their home may affect the distance that an individual can travel from their home and thus impact the spatial scale at which causal processes linking neighborhoods to health may operate. Individual characteristics (e.g., age) may also impact travel patterns. The failure to account for variations in the relevant spatial scale across cities or individuals within cities could result in measurement error in the neighborhood factor, and thus incorrect inference²². To date, however, little work has explored how individual-level characteristics may change the spatial scale of contextual factors.

Using traditional buffer-based approaches, several investigators have conducted empirical studies to examine how the choice of buffer size impacts various environment-health associations and thus help understand a relevant spatial scale on which the environment-health associations operate¹⁹⁻²¹. However, no studies of which we are aware have systematically investigated whether the magnitude of the associations and/or spatial scale vary as a function of individual characteristics, in part due to lack of available methods to do so. Distributed lag models (DLMs) were proposed for examining how the built environment-health association decreases/increases as a function of continuous distance from locations of interest without a priori specification of a buffer size¹⁶. By letting the association vary continuously with distance from residential locations, DLMs allows indirect inference on the

spatial scale and circumvent the need to specify the buffer size a priori. However, DLMs are limited to examining associations at the population level.

Hierarchical distributed lag models (HDLMs), an extension of DLMs for hierarchical data²³, can be used to investigate the between-group and between-individual heterogeneity in the scale and magnitude in the environment-health associations by modeling the distributed lag (DL) coefficients as random. By explicitly modeling individual- and site-specific heterogeneity, HDLMs can be used to examine whether and how the associations between features of neighborhood environments and health vary in the magnitudes or spatial scales across individuals within/across contexts. A previous application of HDLMs examined heterogeneity in associations and spatial scales across administrative areas (assembly districts)²³. Here we use HDLMs to additionally investigate said heterogeneity across individual persons and contexts.

The goal is to illustrate HDLMs as an approach to systematically explore heterogeneity of the spatial scale and magnitude for the neighborhood environment-health association and heterogeneity in scales across individuals, neighborhoods and cities. We use longitudinal data from the Multi-Ethnic Study of Atherosclerosis (MESA) for the time period 2000–2010 and implement HDLMs in a Bayesian framework²³ to: 1) examine whether and how the association between the availability of favorable food stores (defined as supermarkets and fruit and vegetable markets) and body mass index (BMI) depends on continuous distance from participants' locations, thus allowing us to indirectly infer the spatial scale at which the association operates 2) examine if the spatial scale and magnitude of the association differs across six MESA sites and 3) across individuals.

MATERIALS AND METHODS

Study Population

The MESA is a longitudinal study investigating the determinants of subclinical cardiovascular disease in six US cities (Baltimore, MD; Chicago, IL; Forsyth County, NC; Los Angeles, CA; New York, NY; and St. Paul, MN). Objectives and design of the MESA study were previously reported²⁴. The study recruited 6814 participants aged 45 to 84 years at baseline, free of cardiovascular disease. A baseline assessment was conducted in 2000–2002, with four follow-up exams occurring at approximately 1.5–2 year intervals. The MESA neighborhood study, an ancillary study to the MESA, had the overarching goal of studying how features of neighborhoods influence CVD and CVD risk factors²⁵; this ancillary study geocoded residential addresses for all MESA participants who agreed to participate. Of the 6814 MESA participants, 623 subjects did not participate in the neighborhood study. Given the hierarchical nature of the HDLM, repeat visits for participants were included in analyses if at the time of those visits the participants remained in their baseline MESA site. Participants' visits with incomplete information on exposure, outcome or covariates were excluded. Supplementary Table 1 gives the specific number of visits and participants excluded. This study included data from 26,691 study visits from 6,143 unique participants. The study was approved by the Institutional Review Board of all participating institutions and all participants provided informed consent.

Individual Measures

BMI (weight (kg)/(height (m)²) was used as the outcome. Adjustment factors included sex, race/ethnicity, age at exam 1, education; and total annual family income, marital status, smoking status and cancer diagnosis as time varying covariates. Total annual gross family income was categorized in 13 classes, and then a continuous income measure was generated as the interval midpoint of the selected category. Since marital status and total annual gross family income were only available for three out of five exams, the marital status and income values from the previous exam were carried forward.

Neighborhood Measures

Measures of street connectivity and population density were constructed for each participant's residential address. Street connectivity (network ratio) was created as the proportion of a 1-mile Euclidean-distance buffer that is covered by a buffer creating using network distance (range 0–1, with a higher proportion indicating better street connectivity²⁶). Population density (per square mile) in a 1-mile buffer was created based on block-level population data from Census 2000 or 2010. A weighted factor scale for neighborhood overall SES was created by principle factor analysis of 16 census variables, which reflected aspects of education, occupation, household income and wealth, poverty, employment, and housing from Census 2000²⁷ and American Community Surveys 2005–2009²⁸ and 2007–2011²⁹ as described elsewhere³⁰. A higher value of this score reflects lower neighborhood SES.

Food Environment Data

Food establishment data were obtained from the National Establishment Time Series (NETS) database for 2000–2010³¹. Favorable food stores were defined as supermarkets or fruit/vegetable markets based on a total of 15 Standardized Industrial Codes (SIC) along with supermarket names derived from the Nielsen/TDLinx database³². Locations of food stores for every ZIP code within a 5 mile radius of participants' residential locations were obtained. This distance is longer than that used by the USDA to define urban food deserts, which are areas that, even within 1 mile of them, are “vapid of fresh fruit, vegetables, and other healthful whole foods” (the distance is 10 miles for rural areas)³³. Distances between each food store and each participant's residential address were then calculated, and counts of food stores within ring-shaped areas were used as predictor variables as detailed below. Since addresses of food stores around participant's residential locations may change over time, geocodes for participants' residential locations and food stores were cross-referenced for each visit.

Statistical Model

The outcome Y_{ijs} is the BMI of participant i in visit j at MESA site s . Measures of the favorable food environment are denoted $X_{ijs}(r_{l-1}; r_l)$, $l = 1, 2, \dots, L$, which are counts of favorable food stores within ring-shaped areas around the address of the i^{th} participant at visit j , with inner and outer radius r_{l-1} and r_l respectively. The basic DLM used is

$$Y_{ijs} = \beta_{0i} + \beta_{1i} \text{time}_{js} + \sum_{l=1}^L \beta(r_{l-1}; r_l) X_{ijs}(r_{l-1}; r_l) + \mathbf{Z}_{ijs} \boldsymbol{\gamma} + \varepsilon_{ijs}, \quad (1)$$

where $\varepsilon_{ijs} \sim N(0, \tau^2)$ is a residual error; β_{0i} represents a random intercept of subject i , $\beta_{0i} \sim N(\beta_0, \sigma_u^2)$, to account for unobserved time-invariant covariates and centered at the overall mean BMI β_0 ; β_{1i} represents a random time effect of subject i , $\beta_{1i} \sim N(\beta_1, \sigma_v^2)$, to capture between-individual differences in change in BMI and centered at the overall mean rate of change in BMI, β_1 ; the DL coefficient $\beta(r_{l-1}; r_l)$ denotes the association of the food environment measured between radius r_{l-1} and r_l around participant's residential locations; $\mathbf{Z}_{ijs}$ denotes the set of covariates previously described, with corresponding coefficients $\boldsymbol{\gamma}$. We used $r_L = 5$ miles as the maximum distance from the residential locations. The total number of lags $L = 50$ was used to allow for flexibility in the shapes of the associations of the aggregated environment feature and the outcome, and is deemed large enough to provide consistent estimates¹⁶. Thus, the HDLM estimates the associations at every 0.1 of a mile.

Since we would not typically expect associations to change abruptly across distance, we modeled the DL coefficients $\beta(r_{l-1}; r_l)$ a smooth function of distance r_l , $l = 1, 2, \dots, L$, using splines^{34,35}. That is, we constrained the coefficients corresponding to adjacent ring-shaped areas to be similar, by modeling them as:

$$\beta(r_{l-1}; r_l) = \alpha_0 + \alpha_1 r_l + \sum_{k=1}^L \tilde{\alpha}_k |r_l - r_k|^3, \quad (2)$$

where α_0 denotes the intercept of the lag effects, α_1 represents the average change of the association across distance, and $\tilde{\alpha}_k$ are coefficients that capture deviations from linearity and are penalized to achieve smoothness.

To enable (some) comparison with the traditional buffer-based approach, we used the model coefficients to derive average buffer-based associations. For given a-priori buffer radius (i.e., a distance r_k such as 1 mile), we derived the average association within the buffer of radius

r_k as $\bar{\beta}(r_k) = \sum_{l=1}^k \beta(r_{l-1}; r_l) \pi(r_l^2 - r_{l-1}^2) / \pi r_k^2$, an area-weighted average of the coefficients and provides the estimated difference of BMI per additional food store within the buffer of radius r_k . However, in contrast to traditional buffer-based approaches, $\bar{\beta}(r_k)$ is adjusted for the availability of food stores between r_k and r_L , which could confound $\bar{\beta}(r_k)$.

Starting from the basic model in equation (1) (referred below as “Model 1”), we fitted additional models, detailed below, to investigate which additional DL terms explain variation in BMI. We used the deviance information criterion (DIC) to compare models. DIC is a measure of goodness-of-fit in the Bayesian framework that trades off between model fitness and complexity³⁶; a lower DIC value indicates a better model. The details of fitting HDLMs in a Bayesian framework were previously described³⁷, although here the models are extended to include site-specific DL coefficients and individual-specific DL coefficients within sites. R codes used to fit all models are provided in the supplementary materials.

Compared to Model 1, Model 2 includes interactions between the food environment associations defined in equation (2) and MESA sites. For site s , we obtain a site-specific association, $\beta_s(r_{l-1}; r_l) = \alpha_{0s} + \alpha_{1s}r_l + \sum_{k=1}^L \tilde{\alpha}_{ks}|r_l - r_k|^3$. This model captures differences in the DL coefficients across sites, but the coefficients are the same for all locations and participants within sites.

Model 3 investigates individual-level heterogeneity in the DL coefficients by treating them as random. We modeled $\beta_{ijs}(r_{l-1}; r_l) = \alpha_{0is} + \alpha_{1is}r_l + \sum_{k=1}^L \tilde{\alpha}_{ks}|r_l - r_k|^3$, where

$\begin{bmatrix} \alpha_{0is} \\ \alpha_{1is} \end{bmatrix} \sim N\left(\begin{bmatrix} \alpha_{0s} \\ \alpha_{1s} \end{bmatrix}, \Sigma_s\right)$ with unstructured covariance matrix Σ_s for each site s , i.e., the intercept of lag effects and the linear change of the association across distance vary across individuals. If denser sampling of individuals was available, modeling individual non-linear deviations ($\tilde{\alpha}_{ks}$) could be possible, in principle. This model assumes the random effect variances differ by site, i.e., $\Sigma_1 \cdots \Sigma_6$. Fitting a restricted version of Model 3 (Model 3R) assuming $\Sigma_1 = \cdots = \Sigma_6$ enables us to distinguish if the degree of individual heterogeneity of the DL coefficients is the same or different across MESA sites.

Next, we examined which observed individual-level covariates explained the variation of individual DL coefficients (α_{0is} , α_{1is}) in Model 3 by allowing the DL coefficients to depend on observed factors. Specifically, we fitted:

$\beta_{ijs}(r_{l-1}; r_l) = \alpha_{0is} + \alpha_{1is}r_l + \sum_{k=1}^L \tilde{\alpha}_{ks}|r_l - r_k|^3 + (\theta_{0s} + \theta_{1s}r_l)Z_{ijs}$, where Z_{ijs} is the baseline age, sex, race/ethnicity, income, street connectivity, population density, or time since baseline, and where the θ_{0s} and θ_{1s} jointly capture the differences in the intercept and linear decay of the lag effects associated with one unit higher Z_{ijs} in site s . Modeling between-individual differences only in the intercept and slope (θ_{0s} and θ_{1s}) is consistent with Model 3, since Model 3 captures only individual heterogeneity in the DL's intercepts and slope. These models were fitted separately for each variable, but for brevity are called "Model 4".

RESULTS

At baseline, the mean age was 61.8 (standard deviation (SD) = 10.1) years, 52.4 percent were female, 39.3, 11.8, 27.2, and 21.7 percent were White, Chinese American, African American, and Hispanic, respectively, although race/ethnicity was not evenly distributed by site (Table 1). The mean BMI was 28.3 (SD = 5.41) kg/m². The mean number of favorable food stores in buffers from 1 to 5 miles are 7.3 (SD = 11.1), 18.8 (SD = 27), 24.4 (SD = 33), 30.6 (SD = 38.9), and 34.3 (SD = 41.3), respectively. Across the sites, NC and MN had the lowest number of favorable food stores, while NY had the highest availability and also the largest between-individual variability in the available food stores.

Figure 1A–B shows "unadjusted" (except for sites and time since baseline) estimates for the overall DL association between favorable food environment and BMI up to 5 miles and buffer estimates. In this unadjusted model, the spatial scale where food environment-BMI association operates is approximately 3–4 mile. The estimated posterior probability that a buffer association is negative was 0.96 at 4 mile but 0.66 at 5 mile.

After adjusting for individual- and neighborhood variables in Model 1 (all other model coefficients are shown in Supplementary Table 2), the association between more favorable food environments with lower BMI was attenuated but still tended to be stronger at shorter distances between favorable food stores and residential locations, confirming the intuitive notion that smaller spatial scales are more relevant. The point estimate for the association declined quickly up to approximately 2 miles, implying food environment within this distance have larger associations with BMI.

The average association within buffers was much stronger for closer distances (Figure 1D). The estimated average association on BMI for a 1-mile buffer was -0.004 [95% credible interval (CI): $-0.012, 0.003$] per additional food store; the estimated posterior probability that a 1-mile buffer association is negative was 0.87. In contrast, the estimated average association with a 5-mile buffer was 0.0001 (95% CI: $-0.001, 0.002$) per one food store increase; the estimated posterior probability that a 5-mile buffer association is negative was 0.40. The associations at other distances shown in Figure 1B can be found in Supplementary Table 3.

Compared to Model 1 (DIC=113571), Model 2 (DIC = 113577) yielded a slightly higher DIC, suggesting that the DL coefficients capturing the food environment-BMI association across sites were not different. However, Model 3 had a lower (DIC = 134422) indicating that there is individual-level heterogeneity in the DL coefficients, and comparing Model 3 to a restricted version (DIC Model 3R= 134578) indicated that the degree of individual-level heterogeneity was the different across sites. The results from Model 3 thus support the idea that site- and individual-level differences in DL coefficients exist.

Figure 2 shows the average DL coefficients within sites, while Supplementary Figure S1 demonstrates the degree of heterogeneity in the individual-level associations in each site. In NY, IL and MN sites the DL coefficients were negative at shorter distances, although the NY estimates had higher precision (Figure 2). The similarities in the pattern of the DL coefficients across distance from residential locations in these three sites suggests that the spatial scales in these sites are likely to be similar. The DL coefficient estimates in MD, CA and NC were either zero or positive, although with a large degree of uncertainty. The individual-level variation in the DL coefficients was largest in NC and smallest in NY and CA (see Supplementary Figure S1).

Using Model 4 we examined if the individual-level heterogeneity could be explained by individual characteristics (baseline age, sex, race/ethnicity, income), characteristics of the individual's residential location (street connectivity or population density), or time since baseline (e.g., aging). The only characteristic that resulted in a lower DIC was sex (DIC =113431). DIC values from all the models can be found in Supplementary Table 4.

Figure 3 shows the point estimates of the associations of the food environment with BMI by sex and site for distances up to 5 miles from residential locations derived from Model 4. Sex differences in the food environment-BMI association were more pronounced in NC and MN than the other sites. The apparent lack of food environment-BMI association in NC (Figure 2) is likely a consequence of the positive association among men canceling the negative

association among women. In MN, the favorable food environment is more negatively associated with BMI for women at shorter distances from residential locations, indicating the spatial scale for women may be smaller than for men. Estimates for site-specific average (buffer) associations by sex at 1–5 mile show similar trends as in Figure 3 (see Supplementary Figure S2).

DISCUSSION

This study illustrates the use of HDLMs as a tool to examine the built environment-health association as a continuous function of distance from locations of interest, thereby indirectly informing the spatial scale of the association, and to examine variation in the associations or spatial scale across individuals and contexts. We investigated heterogeneity of spatial scale and magnitude in the favorable food environment-BMI association. We found some evidence supporting the idea that there is between-site and between-individual heterogeneity in the spatial scale and magnitude in said association. Varying degrees of heterogeneity across sites suggested that contextual factors within sites or individual's characteristics may modify the food environment-BMI association. In particular, we found that sex partially explained individual-level heterogeneity in the spatial scale and magnitude in the food environment-BMI association.

A growing literature has begun to examine the extent of potential biases introduced by an arbitrary choice of spatial scale^{15,19,38} and approaches to identify the relevant spatial scales at which the environment-health association operates^{20,21}. Our analysis allowed us to closely examine the way that the relationship between individual's neighborhood resources and BMI varied as a function of continuous distance in a longitudinal cohort study. While the directions of our findings were consistent with previous MESA studies^{6,39}, our investigation provides evidence that stronger inverse associations between the food environment and BMI occur at shorter distances between the stores and residential locations. It is plausible that presence of stores closer to the home is more relevant to BMI because individuals may be more likely to access stores closer to home due to convenience, ease, or transportation access⁴⁰. It is important to note that unadjusted models illustrated a relevant distance of approximately 3–4 miles, while models adjusting for relevant covariates identified the shorter distance of about 2 miles. This may be due to the fact that the covariates included are also spatially patterned. Hence, the adjusted distance represents a more direct estimate of the distance of this association. Even in the absence of concrete evidence of the causes behind this pattern this result can have implications for explaining inconsistencies found across other studies; it is plausible that research projects using smaller buffers would find a stronger association than those which use large buffers or administrative units (although both with potential biases as noted by Baek et al.¹⁶). More generally identifying the spatial scale potentially involved in causal effects of environments on outcomes has important implications for interventions¹².

In addition, HDLMs allowed us to examine between-site and between-individual heterogeneity in the spatial scale and magnitude for the favorable food environment-BMI association. We found that there is between-individual heterogeneity and the degree of between-individual heterogeneity was different across sites. Among the observed factors,

only sex explained between-individual heterogeneity in the spatial scale and magnitude in the food environment-BMI association across sites. For instance, in NC and MN, the food environment-BMI association was more inversely related for women compared to men at a shorter scale from residential locations. Although not directly comparable, previous research reported a similar finding of the food environment and sex interaction⁴¹. To the extent that variability in the scale and magnitude of in associations may be due to unobserved differences in individuals' activity spaces driven by macro-level contextual factors, our results suggest that NY participants have activity spaces more similar to each other compared to individuals in NC. Several studies have begun to use novel approaches, such as Global Positioning Systems (GPS) to understand individual activity spaces^{42–44}. Another possibility is that said variability results from differences in urban form, including localized transportation resources to access food stores, which were unmeasured. Alternately, individuals living in different environments may be conditioned to expect or anticipate different travel time. While these new methods attempt to accommodate and account for the uncertainty due to unobserved individual characteristics by sites, few studies have examined explanations for between-individual heterogeneity⁴³. Our approach allowed us to examine and quantify this heterogeneity.

As a result of the observational design of our study, we were unable to make causal inferences about the relationship between availability of favorable food stores and BMI. The circular shape of the neighborhoods may not represent the true underlying irregular shape of individuals' activity spaces, introducing measurement error into our exposure measures⁴⁵. Although we adjusted for neighborhood SES, street connectivity, and population density, there may be residual confounding due to other neighborhood features. We did not adjust for mediators, such as diet or intentional physical activity⁶, since we sought to measure the total association between favorable food environment and BMI. While our study examined heterogeneity in the association by individual factors, an older adult sample may have limited our ability to examine heterogeneity across a broader age spectrum. Similarly, our examination of heterogeneity by city was limited to the six MESA sites, which may not represent smaller cities or rural areas in other geographic regions. Finally, while our application of HDLMs represents a case study to test an emergent methodology, this paper only sheds light on variations in the neighborhood food environment-BMI relationship. Future work should apply HDLMs to research examining other neighborhood features and health indicators.

This is the first study applying HDLMs to cohort data for examining the food environment-BMI association. HDLMs allowed us to examine the food environment-BMI association across a large spectrum of distances from individual's residential locations, thereby indirectly inferring the spatial scale for the association, and study the variations of the spatial scale and association across individuals and larger contexts. In contrast to other studies we did not specify buffers or areas a priori. Rather than ignoring or mentioning as limitations the potential differences in spatial scales across individuals and contexts, future work should consider employing methodologies, such as HDLMs, that allow more nuanced investigation of various spatial scales in a systematic manner. This may enhance our ability to build theories about the relevant spatial scales for environmental effects that can subsequently be further examined and tested using a variety of qualitative and quantitative approaches.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 from the National Heart, Lung, and Blood Institute and by grants UL1-TR-000040, UL1-TR-001079, UL1-RR-024156 and UL1-RR-025005 from NCRR, R01 HL071759 (PI: Diez Roux) from National Heart, Lung, and Blood Institute at the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. This research received support from the Population Research Training grant (T32 HD007168) and the Population Research Infrastructure Program (R24 HD050924) awarded to the Carolina Population Center at The University of North Carolina at Chapel Hill by the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. A full list of participating MESA investigators and institutions can be found at <http://www.mesa-nhlbi.org>.

REFERENCES

- Black JL, Macinko J. Neighborhoods and obesity. *Nutr Rev*. 2008; 66(1):2–20. [PubMed: 18254880]
- Ludwig J, Sanbonmatsu L, Gennetian L, et al. Neighborhoods, Obesity, and Diabetes — A Randomized Social Experiment. *N Engl J Med*. 2011; 365(16):1509–1519. [PubMed: 22010917]
- Morland K, Diez Roux AV, Wing S. Supermarkets, other food stores, and obesity: The Atherosclerosis Risk in Communities Study. *Am J Prev Med*. 2006; 30(4):333–339. [PubMed: 16530621]
- Diez-Roux AV. Bringing Context Back into Epidemiology: Variables and Fallacies in Multilevel Analysis. *Am J Public Health*. 1998; 88(2):216–222. [PubMed: 9491010]
- Susser M. The logic in ecological: I. The logic of analysis. *Am J Public Health*. 1994; 84(5):825–829. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1615050&tool=pmcentrez&rendertype=abstract>. [PubMed: 8179056]
- Mujahid MS, Roux AVD, Shen M, et al. Relation between neighborhood environments and obesity in the multi-ethnic study of atherosclerosis. *Am J Epidemiol*. 2008; 167(11):1349–1357. [PubMed: 18367469]
- Moore LV, Diez Roux AV, Nettleton JA, Jacobs DR. Associations of the local food environment with diet quality - A comparison of assessments based on surveys and geographic information systems. *Am J Epidemiol*. 2008; 167(8):917–924. [PubMed: 18304960]
- Moore LV, Diez Roux AV, Nettleton JA, Jacobs DR, Franco M. Fast-Food consumption, diet quality, and neighborhood exposure to fast food. *Am J Epidemiol*. 2009; 170(1):29–36. [PubMed: 19429879]
- Rundle A, Neckerman KM, Freeman L, et al. Neighborhood food environment and walkability predict obesity in New York City. *Environ Health Perspect*. 2009; 117(3):442–447. [PubMed: 19337520]
- Bodur JN, Rice JC, Farley TA, Swalm CM, Rose D. The association between obesity and urban food environments. *J Urban Heal*. 2010; 87(5):771–781.
- Inagami S, Cohen DA, Brown AF, Asch SM. Body mass index, neighborhood fast food and restaurant concentration, and car ownership. *J Urban Heal*. 2009; 86(5):683–695.
- Diez-Roux AV. Investigating neighborhood and area effects on health. *Am J Public Health*. 2001; 91(11):1783–1789. [PubMed: 11684601]
- Papas MA, Alberg AJ, Ewing R, Helzlsouer KJ, Gary TL, Klassen AC. The built environment and obesity. *Epidemiol Rev*. 2007; 29(1):129–143. [PubMed: 17533172]
- Diez Roux AV, Mair C. Neighborhoods and health. *Ann N Y Acad Sci*. 2010; 1186(1):125–145. [PubMed: 20201871]

15. Spielman SE, Yoo E-H. The spatial dimensions of neighborhood effects. *Soc Sci Med.* 2009; 68(6): 1098–1105. [PubMed: 19167802]
16. Baek J, Sánchez BN, Berrocal VJ, Sanchez-vaznaugh EV. Distributed Lag Models: Examining Association Between the Built Environment and Health. *Epidemiology.* 2016; 27(1):116–124. [PubMed: 26414942]
17. Fotheringham, aS, Wong, DWS. The modifiable areal unit problem in multivariate statistical analysis. *Environ Plan A.* 1991; 23(7):1025–1044.
18. Openshaw, S. Chapter 4. Developing GIS-Relevant Zone-Based Spatial Analysis Methods. In *Spatial Analysis: Modelling in a GIS Environment*, Edited by P. Longley and M. Batty. Batty, M., Longley, P., editors. Cambridge, England: Geoinformation International; 1996.
19. James P, Berrigan D, Hart JE, et al. Effects of buffer size and shape on associations between the built environment and energy balance. *Health Place.* 2014; 27:162–170. [PubMed: 24607875]
20. Chaix B, Merlo J, Subramanian SV, Lynch J, Chauvin P. Comparison of a Spatial Perspective with the Multilevel Analytical Approach in Neighborhood Studies : The Case of Mental and Behavioral Disorders due to Psychoactive Substance Use in Malmo, Sweden, 2001. *Am J Epidemiol.* 2005; 162(2):171–182. [PubMed: 15972939]
21. Leal C, Bean K, Thomas F, Chaix B. Are associations between neighborhood socioeconomic characteristics and body mass index or waist circumference based on model extrapolations? *Epidemiology.* 2011; 22(5):694–703. [PubMed: 21709560]
22. Kwan M-P. The Uncertain Geographic Context Problem. *Ann Assoc Am Geogr.* 2012; 102(5):958–968.
23. Baek J, Sanchez-vaznaugh EV, Sánchez BN. Hierarchical Distributed-Lag Models : Exploring Varying Geographic Scale and Magnitude in Associations Between the Built Environment and Health. *Am J Epidemiol.* 2016; 183(6):583–592. [PubMed: 26888753]
24. Bild DE, Bluemke Da, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: Objectives and design. *Am J Epidemiol.* 2002; 156(9):871–881. [PubMed: 12397006]
25. Mujahid MS, Diez-Roux AV, Morenoff JD, Raghunathan T. Assessing the measurement properties of neighborhood scales: From psychometrics to econometrics. *Am J Epidemiol.* 2007; 165(8):858–867. [PubMed: 17329713]
26. Hirsch JA, Moore KA, Barrientos-Gutierrez T, et al. Built environment change and change in BMI and waist circumference: Multi-ethnic Study of Atherosclerosis. *Obesity.* 2014; 22(11):2450–2457. [PubMed: 25136965]
27. Bureau of the Census, US Department of Commerce. Year 2000 STF1 and STF3. Washington DC: Bureau of the Census; 2000.
28. Bureau of the Census, US Department of Commerce. American Community Survey 5-year small area estimates 2005–2009. Washington DC: Bureau of the Census; 2010.
29. Bureau of the Census, US Department of Commerce. American Community Survey 5-year small area estimates 2007–2011. Washington DC: Bureau of the Census; 2012.
30. Moore K, Roux AVD, Auchincloss A, et al. Home and Work Neighborhood Environments in Relation to Body Mass Index: The Multi-Ethnic Study of Atherosclerosis (MESA). *J Epidemiol Community Heal.* 2013; 67(10):846–853.
31. Business Dynamics Research Consortium. National Establishment Time-Series (NETS) Database. Denver, Colorado: Walls & Associate; 2015. Available at: <http://exceptionalgrowth.org/our-databases.iegc> [Accessed April 5, 2015]
32. Auchincloss AH, Moore KaB, Moore LV, Diez Roux AV. Improving retrospective characterization of the food environment for a large region in the United States during a historic time period. *Health Place.* 2012; 18(6):1341–1347. [PubMed: 22883050]
33. Gallagher M. USDA Defines Food Deserts. *Am Nutr Assoc.* 2010 [Accessed May 18, 2016] Available at: <http://americannutritionassociation.org/newsletter/usda-defines-food-deserts>.
34. Hastie, TJ., Tibshirani, RJ. Generalized Additive Models. Boca Raton, FL: CRC Press; 1990.
35. Zanobetti A, Wand MP, Schwartz J, Ryan LM. Generalized additive distributed lag models : quantifying. *Biostatistics.* 2000; 1(3):279–292. [PubMed: 12933509]
36. Spiegelhalter DJ, Best NG, Carlin BP, Linde A Van Der. Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol.* 2002; 64(4):583–639.

37. Baek J, Sánchez BN, Sanchez-Vaznaugh EV. Hierarchical Distributed Lag Models: Exploring Varying Geographic Scale and Magnitude in Health-Built Environment Associations. *Am J Epidemiol.* 2015 in press.
38. Vallée J, Le Roux G, Chaix B, Kestens Y, Chauvin P. The “constant size neighbourhood trap” in accessibility and health studies. *Urban Stud.* 2014;1–20.
39. Auchincloss AH, Mujahid MS, Shen M, Michos ED, Whitt-Glover MC, Diez Roux AV. Neighborhood health-promoting resources and obesity risk (the multi-ethnic study of atherosclerosis). *Obesity.* 2013; 21(3):621–628. [PubMed: 23592671]
40. Hirsch JA, Hillier A. Exploring the role of the food environment on food shopping patterns in Philadelphia, PA, USA: a semiquantitative comparison of two matched neighborhood groups. *Int J Environ Res Public Health.* 2013; 10(1):295–313. [PubMed: 23343984]
41. Frank L, Kerr J, Saelens B, Sallis J, Glanz K, Chapman J. Food outlet visits, physical activity and body weight: variations by gender and race-ethnicity. *Br J Sports Med.* 2009; 43(2):124–131. doi: [PubMed: 19042921]
42. Burgoine T, Jones AP, Namenek Brouwer RJ, Benjamin Neelon SE. Associations between BMI and home, school and route environmental exposures estimated using GPS and GIS: do we see evidence of selective daily mobility bias in children? *Int J Health Geogr.* 2015; 14(8):1–13. [PubMed: 25563056]
43. Hirsch JA, Winters M, Clarke P, McKay H. Generating GPS activity spaces that shed light upon the mobility habits of older adults: a descriptive analysis. *Int J Health Geogr.* 2014; 13(1):1–14. [PubMed: 24383521]
44. Hirsch JA, Winters M, Ashe MC, Clarke PJ, McKay HA. Destinations That Older Adults Experience Within Their GPS Activity Spaces: Relation to Objectively Measured Physical Activity. *Environ Behav.* 2015:1–23.
45. Chaix B, Merlo J, Evans D, Leal C, Havard S. Neighbourhoods in eco-epidemiologic research : Delimiting personal exposure areas. A response to Riva, Gauvin, Apparicio and Brodeur. *Soc Sci Med.* 2009; 69(9):1306–1310. [PubMed: 19692161]

Figure 1.

A) Sites and time adjusted estimates for overall distributed lag associations with 95 credible interval (CI) between favorable food environment and BMI over distances up to 5 miles. B) Sites and time adjusted estimates for overall average association (buffer association) with 95% CI between favorable food environment and BMI at 1–5 miles. C) Adjusted estimates for overall distributed lag associations with 95 credible interval (CI) between favorable food environment and BMI over distances up to 5 miles. D) Adjusted estimates for overall

average association (buffer association) with 95% CI between favorable food environment and BMI at 1–5 miles. Adjusted estimates are from Model 1.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2.

Estimates for the site-specific distributed lag coefficients (continuous black lines) with 95% CI (dashed line) capturing the association between favorable food environment and BMI over distances up to 5 miles, and estimates for the average association (buffer association) with 95% CI at 1–5 miles are overlaid (vertical lines at 1,..., 5 mile; blue). MESA sites: Baltimore, MD; Chicago, IL; Forsyth County, NC; Los Angeles, CA; New York, NY; and St. Paul, MN. Estimates are from Model 3.

Figure 3.

Sex- and site-specific estimates of the distributed lag associations between favorable food environment and BMI over distances up to 5 miles. MESA sites: Baltimore, MD; Chicago, IL; Forsyth County, NC; Los Angeles, CA; New York, NY; and St. Paul, MN. Estimates are from Model 4 with distributed lag coefficients varied by sex.

Table 1

Descriptive statistics of participants included in the analysis at baseline (Exam 1). Multi-Ethnic Study of Atherosclerosis, 2000–2002 (N=6143).

Variable	level	Overall					MESA site				
		Mean (SD) or %	NC	NY	MD	MN	IL	CA			
N		6143	973	1009	944	969	1072	1176			
BMI		28.3 (5.4)	28.8 (5.6)	28.8 (5.5)	29.4 (5.6)	29.5 (5.3)	26.7 (5.0)	27.1 (5.1)			
Demographic											
Age (years)		61.9 (10.1)	62.2 (9.7)	61.5 (10.1)	63.0 (10.0)	60.1 (10.4)	62.0 (9.9)	62.4 (10.4)			
Female		52.4	52.9	56.1	52.3	50.5	53.0	49.9			
Race/ethnicity	White, Caucasian	39.3	54.8	20.6	49.7	58.3	48.1	10.4			
	Chinese American	11.8	0.0	0.2	0.0	0.0	26.4	37.6			
	Black, African-American	27.2	44.9	34.4	50.3	0.0	25.5	11.6			
	Hispanic	21.7	0.3	44.8	0.0	41.7	0.0	40.4			
Individual characteristics											
Education	HS/GED or less	34.9	28.8	43.9	30.0	39.8	14.5	50.9			
	Some college, Technical or Associate degree	28.4	30.3	27.5	29.7	35.1	24.3	24.9			
	Bachelor's degree or higher	36.7	40.9	28.6	40.4	25.1	61.3	24.2			
Income per \$1000		49.6 (34.1)	56.3 (32.7)	42.5 (29.9)	51.2 (32.9)	44.1 (29.5)	69.1 (37.6)	35.5 (30.0)			
Currently married/living with a partner		61.7	69.1	51.0	56.6	60.0	64.3	67.7			
Cancer		7.9	12.5	6.2	8.6	8.2	8.0	4.7			
Smoking status	Never	50.4	45.1	50.7	46.4	43.6	51.8	62.1			
	Former	36.9	41.5	35.3	40.4	40.5	37.6	28.1			
	Current	12.7	13.4	14.0	13.2	16.0	10.6	9.8			
Environmental characteristics											

Variable	level	Overall	MESA site					
			NC	NY	MD	MN	IL	CA
			Mean (SD) or %	Mean (SD) or %	Mean (SD) or %	Mean (SD) or %	Mean (SD) or %	Mean (SD) or %
Street Connectivity		0.42(0.2)	0.23(0.1)	0.5(0.1)	0.4(0.2)	0.45(0.1)	0.47(0.1)	0.47(0.1)
Population density per 1000 in 1 mile		15.6(19.3)	1.6(0.9)	55.4(14.0)	6.9(4.5)	4.7(1.6)	13.9(6.0)	10.7(4.7)
Favorable food stores up to								
	1 mile	7.3(11.1)	1.1(1.3)	29.4(11.0)	2.6(3.7)	1.5(0.9)	4.9(3.2)	4.4(3.1)
	2 mile	18.8(27.0)	3.2(2.3)	74.9(18.5)	7.0(6.4)	4.4(1.7)	9.3(5.5)	13.3(12.1)
	3 mile	24.4(33.0)	4.7(3.2)	94.4(17.3)	9.2(5.7)	5.2(2.3)	12.0(6.1)	19.7(14.1)
	4 mile	30.6(38.9)	5.8(3.6)	110.6(26.8)	11.1(6.8)	8.4(2.9)	17.6(7.6)	28.3(17.0)
	5 mile	34.3(41.3)	7.0(3.8)	116.1(34.9)	13.3(7.3)	9.5(2.7)	21.5(9.0)	35.8(19.3)