Published in final edited form as:
Ann Epidemiol. 2021 September ; 61: 1–7. doi:10.1016/j.annepidem.2021.05.005.

Comparing competing geospatial measures to capture the relationship between the neighborhood food environment and diet

Pasquale E. Rummo, PhD, MPH[1], **Yasemin Algur, MPH**[2], **Tara McAlexander, PhD**[2], **Suzanne E. Judd, PhD**[3], **Priscilla M. Lopez, MPH**[1], **Samrachana Adhikari, PhD**[1], **Janene Brown, MPH**[2], **Melissa Meeker**[2], **Leslie A. McClure, PhD, MS**[2], **Brian Elbel, PhD**[1,4]

[1]Department of Population Health, New York University Grossman School of Medicine, New York, NY

[2]Dornsife School of Public Health, Drexel University, Philadelphia, PA

[3]Univeristy of Alabama at Birmingham, Birmingham, AL

[4]Wagner Graduate School of Public Service, New York University, New York, NY

Abstract

Purpose: To examine how the choice of neighborhood food environment definition impacts the association with diet.

Methods: Using food frequency questionnaire data from the Reasons for Geographic and Racial Differences in Stroke study at baseline (2003–2007), we calculated participants' dietary inflammation score (DIS) (n=20,331); higher scores indicate greater pro-inflammatory exposure. We characterized availability of supermarkets and fast food restaurants using several geospatial measures, including density (i.e., counts/km^2) and relative measures (i.e., percentage of all food stores or restaurants); and various buffer distances, including administrative units (census tract) and empirically-derived buffers ("classic" network, "sausage" network) tailored to community type (higher-density urban, lower-density urban, suburban/small town, rural). Using generalized estimating equations, we estimated the association between each geospatial measure and DIS, controlling for individual- and neighborhood-level sociodemographics.

Corresponding author: Pasquale E. Rummo, PhD, MPH, pasquale.rummo@nyulangone.org, Full postal address: 180 Madison Avenue, 3rd Floor, #3-54, New York, NY 10016.

AUTHOR CONTRIBUTIONS

Pasquale E. Rummo: Conceptualization; Formal analysis; Methodology; Visualization; Writing - original draft; Writing - review & editing
Yasemin Algur: Conceptualization; Data curation; Formal analysis; Methodology; Resources; Writing - original draft; Writing - review & editing
Tara McAlexander: Conceptualization; Data curation; Formal analysis; Methodology; Resources; Writing - review & editing
Suzanne Judd: Conceptualization; Investigation; Writing - review & editing
Priscilla Lopez: Writing - review & editing
Samrachana Adhikari: Writing - review & editing
Janene Brown: Resources; Writing - review & editing
Melissa Meeker: Resources; Writing - review & editing
Leslie McClure: Conceptualization; Funding acquisition; Investigation; Methodology; Resources; Supervision; Writing - review & editing
Brian Elbel: Conceptualization; Funding acquisition; Investigation; Methodology; Supervision; Writing - review & editing

Results: The choice of buffer-based measure did not change the direction or magnitude of associations with DIS. Effect estimates derived from administrative units were smaller than those derived from tailored empirically-derived buffer measures. Substantively, a 10% increase in the percentage of fast food restaurants using a “classic” network buffer was associated with a 6.3 (SE=1.17) point higher DIS ($p<0.001$). The relationship between the percentage of supermarkets and DIS, however, was null. We observed high correlation coefficients between buffer-based density measures of supermarkets and fast food restaurants ($r=0.73$–0.83), which made it difficult to estimate independent associations by food outlet type.

Conclusions: Researchers should tailor buffer-based measures to community type in future studies, and carefully consider the theoretical and statistical implications for choosing relative (vs. absolute) measures.

Keywords

Diet; Restaurants; Inflammation

INTRODUCTION

Epidemiologic research using geospatial methods to explain the relationship between the food environment and population health outcomes has grown dramatically over the past several decades. The food environment, including access to food establishments selling a mix of healthy and less healthy food items (e.g., supermarkets) and those selling predominately less healthy food items (e.g., fast food restaurants), is an important modifiable risk factor influencing diet behaviors.(1) Though findings are mixed, previous research suggests that differences in the food environment are linked to variation in diet quality, obesity, diabetes, and other chronic diseases.(2, 3)

One challenge of characterizing the food environment is the modifiable areal unit problem, or measurement error due to variation in the definition of areal units.(4) For example, previous studies have defined the food environment using administrative units (e.g., census tracts) or empirically-derived units (e.g., buffers around a point location), but often do not provide a rationale for their choice.(5) Previous studies also differ in their choice of buffer size and their specification of buffer shapes around point locations, including the operationalization of network-based buffers. These differences in defining the community context may partially explain the heterogeneity of results in the food environment literature.

To accurately assess the impact of the food environment on health outcomes, the choice of measure must be grounded in plausible theory (e.g., how far people typically travel to their primary food destination). However, we know little about how the operationalization of food environment measures influences observed associations with health outcomes. A comparison of measures tailored to a person’s actual behavior versus “one-size-fits-all” measures is lacking.(6) A comparison of such measures is critical because studies lacking tailored buffer measures typically default to fixed geographic units, despite the arbitrariness of administrative boundaries. Previous work also notes how models with relative measures (e.g., proportions) have better model fit, and larger effect sizes, than models with absolute measures.(7, 8) Therefore, it is also important to consider whether the classification of relative and absolute measures influences associations with health outcomes, especially in a geographically-diverse sample.

To fill gaps in the knowledge base, we propose to thoughtfully examine how the choice of geospatial measure influences the relationships between the food environment, diet, and diabetes status among participants in the REasons for Geographic and Racial Differences in Stroke (REGARDS) study. To accomplish these goals, we constructed competing geospatial measures of the food environment using several measures, including administrative units and empirically-derived buffer measures. We also compared different specifications of these geospatial measures, including absolute versus relative measures, and measures using tailored versus “one-size-fits-all” buffer sizes.

MATERIALS AND METHODS

Study sample and procedures

REGARDS is a prospective national population-based cohort study of regional and racial disparities in stroke incidence and mortality in the continental United States.(9) Participants included 30,239 black and white adults aged ≥45 years enrolled between January 2003 and October 2007. The study oversampled Black individuals (42%) and residents in the Stroke Belt region in the southeastern United States (56%). Following enrollment and verbal consent, REGARDS staff collected comprehensive information on medical history by computer-assisted telephone interviewing. At an in-home visit, trained study personnel collected physical measurements, medication inventory, a resting ECG, and blood and urine samples using standard methods.

Dietary data were collected using a self-administered, validated Block-98 food frequency questionnaire (FFQ).(9–11) The FFQ included 109 food items and assessed usual frequency and quantity of each food item consumed over the previous year. Frequency and quantity were multiplied to calculate the amount consumed for each food item by NutritionQuest (Berkeley, CA). In this study, we excluded REGARDS participants missing diet data (n=9,643) or a census tract ID at baseline (n=287), resulting in a sample size of 20,331 participants. Those missing dietary data were more likely to report Black, lower income, and less than a high school education.

Outcomes

Our selection of outcomes was informed by our participation in the Diabetes LEAD (Location, Environmental Attributes, and Disparities) Network.(12) The Diabetes LEAD Network is a CDC-funded collaboration among Drexel University, Geisinger-Johns Hopkins University, New York University Grossman School of Medicine, and University of Alabama at Birmingham. The primary goal of the network is to examine geographic disparities in diabetes incidence across the U.S. The network has identified several potentially modifiable community domains to prioritize, including the socioeconomic, food, and physical activity environments.

Based on these goals, we selected dietary inflammation score (DIS) as our primary outcome. DIS is a validated measure of the contribution of diet to inflammation,(13) and thus plays a key role in diabetes prevention and control.(14) To calculate DIS, responses to the FFQ were aggregated to 19 food groups, and weights were assigned to each food group by assessing associations with inflammatory markers and then summed. Higher scores indicate greater exposure to foods contributing to inflammation (theoretical range: −14.9–12.8). To explore the robustness of our findings, we examined Mediterranean diet score and diabetes status as secondary outcomes. We calculated the former using previously published methods,(15) with a higher score reflecting higher adherence to a Mediterranean diet (theoretical range: 0 –9). We defined diabetes as a fasting serum glucose ≥126 mg/dL (≥7 mmol/L), current use of insulin or oral diabetes medications, or a self-reported physician diagnosis of diabetes while not pregnant at baseline.

Covariates

Covariates included individual-level age, sex, education (less than high school; high school graduate; some college; college graduate and above), race (Black, White), household income (<$20,000; $20,000-$34,000; $35,000-$74,000; ≥$74,000); and neighborhood socioeconomic environment (NSEE). NSEE was defined as a z-score sum of six Census variables, modeled on previous work and scaled to be between 0 and 100.(16) To classify community type, we used a measure developed by Diabetes LEAD Network,(17) which combined Rural Urban Commuting Area codes and land area of participants' residential census tract to create a four-level variable representing higher-density urban, lower-density urban, suburban/small town, and rural communities.

Exposures

We used food establishment data from the Retail Environment and Cardiovascular Disease (RECVD) study,(18) which classified neighborhood amenities using the National Establishment Time Series (NETS) Database. The NETS data were licensed from Walls & Associates (Walls & Associates, Denver, CO), who prepared annual establishment information collected by Dun and Bradstreet (D&B, Short Hills, NJ). The RECVD team re-geocoded the NETS data to improve locational accuracy and assigned establishments to subcategories using Standard Industrial Classification codes, employee and sales information, and chain names obtained from Technomic/Restaurants and Institutions (R&I) and TDLinx®. Details on classification methods have been described elsewhere.(19)

Our exposures included the availability of supermarkets and fast food restaurants. For each category, we used two different geospatial definitions of "availability", including an absolute measure of the density (count/km^2) of supermarkets and fast food restaurants; and a relative measure of the percentage of supermarkets out of total food stores, and the percentage of fast food restaurants out of total restaurants. The supermarkets category included three mutually-exclusive subcategories: supermarkets, supercenters, and medium-sized grocers. Fast food restaurants were defined as quick-service restaurants offering low-preparation-time foods for take-away or cafeterias (no wait service).

We operationalized the absolute and relative measures using three geospatial definitions (Table 1). First, we calculated the availability of food outlets in a participant's census tract, based on 2010 US Census boundaries.(20) Second, we calculated a "classic" network buffer around the population-weighted centroid of the census tracts of participants' home addresses. Third, we calculated a "sausage" network buffer around participants' exact addresses. In addition to using participants' exact addresses, the "sausage" buffer is distinct from the "classic" buffer in that it buffers the street network by a uniform radius of 150 meters from the street centerline.(21) Street network data was obtained from ESRI's ArcGIS StreetMap Premium. The "classic" and "sausage" network buffers were created using the "generalized" polygon option and default settings in ArcGIS Pro 2.4.2 and ArcGIS Pro 2.1, respectively.(30)

We tailored "classic" and "sausage" network buffer measures to the community type of participants' residential census tract in our primary analyses. The Diabetes LEAD Network based their buffer distances for "classic" network buffers on data from the National Household Food Acquisition and Purchase Survey (FoodAPS),(22) which calculated the average driving distance between participants' residential addresses and their primary food store. The FoodAPS data also assigns participants to rural (yes/no) and non-metro (yes/no) categories, which align with our four-level community-type variable. Based on the FoodAPS mean distances within rural and non-metro categories, we assigned 1-, 2-, 6-, and 10-mile buffer distances to participants residing in high-density urban, low-density urban, suburban/small town, rural census tracts, respectively. Researchers who created our "sausage" network buffers based buffer distances on previous studies,(23–27) which allowed us to assign 1-, 5-, 8-, and 16-km buffer distances to participants residing in high-density urban, low-density urban, suburban/small town, rural census tracts, respectively.

Statistical analysis

We summarized characteristics of the REGARDS participants in our sample using chi-square tests for categorical variables and ANOVA for continuous variables. We calculated Pearson's correlation coefficients between REGARDS participants' different food environment measures. To examine the relationship between our food environment measures tailored to community type and DIS, we used a generalized estimating equation with an identity link, equal-correlation structure, robust standard errors, and clustering at the census tract level, controlling for individual-level covariates and NSEE. Supermarkets and fast food restaurants were included in separate regression models. Because proportion measures have been criticized for not addressing the quantity of food stores,(8, 28) we controlled for total food outlets (continuous) in models with relative measures. We used identical approaches to examine the associations between our food environment measures and secondary outcomes, including Mediterranean Diet score (identity link) and diabetes status (logit link). We conducted sensitivity analyses using "one-size-fits-all" buffer sizes for each network buffer measure (e.g., 1-mile buffer size for all participants). We used Stata version 15.1 (StataCorp LP, College Station, TX) for all analyses. All significance tests were two-sided at unadjusted alpha of 0.05.

RESULTS

The mean DIS and Mediterranean Diet score for the overall sample was −0.004 (SD=2.52) and 4.4 (SD=1.7), respectively, and 21% of participants were classified as having diabetes at baseline (Table 2). We observed a high correlation between our empirically-derived density measures (count/km^2) of supermarkets and fast food restaurants (r=0.73 to 0.83); whereas, the correlation between census tract density measures of supermarkets and fast food restaurants was moderate (r=0.33) (Appendix A). The correlation between all relative measures was small to moderate (r=−0.27 to 0.39).

Approximately 16.7%, 39.8%, 19.3%, 24.2% of participants resided in higher-density urban, lower-density urban, suburban/small town, and rural communities, respectively. The density of fast food restaurants was smaller on average using a “one-size-fits-all” census tract measure [0.55 (SD=1.19)] compared to “classic” [0.69 (SD=0.90)] and “sausage” [0.66 (SD=0.95)] buffer measures tailored to community type (Table 3). We also observed smaller values for supermarket density using a “one-size-fits-all” census tract measure versus tailored network buffer measures, and smaller values for relative “one-size-fits-all” census tract measures compared to tailored relative network buffer measures.

In adjusted analyses, we found that a one-unit increase in the density of supermarkets using a “classic” network buffer tailored to community type was associated with a 0.30 (SE=0.05) decrease in DIS (p<0.001) (Table 4). The relationship between the percentage of supermarkets and DIS, however, was not statistically significant. Contrary to our expectations, we found that a one-unit increase in the density of fast food restaurants in a “classic” network buffer tailored to community type was associated with a 0.15 (SE=0.02) **decrease** in DIS (p<0.001). Whereas, a 10% increase in the percentage of fast food restaurants using a “classic” network buffer tailored to community type was associated with a 6.3 (SE=1.17; p<0.001) point **increase** in DIS.

Differences in the effect estimates between “classic” and “sausage” network buffers were negligible across all models. The effect estimates for the “one-size-fits-all” census tract-based measures were smaller than estimates derived from the buffer-based measures tailored to community type; and estimates for the absolute and relative measures tailored to community type were the most similar to estimates derived from the 5-km and 8-km buffer-based measures, respectively (Table 3). In sensitivity analyses, we observed an increase in effect size as the size of the buffer increased. For example, the effect size for the percentage of fast food restaurants increased from 0.26 (SE=0.08; p=0.002) to 0.50 (SE=0.13; p<0.001) to 1.02 (SE=0.19; p<0.001) using a 5-km, 8-km, and 16-km buffer size, respectively.

In secondary analyses, we observed positive associations between the density of supermarkets and fast food restaurants with Mediterranean Diet score, regardless of geospatial definition (Appendix B); and negative associations with diabetes status (Appendix C). The associations with our relative measures of supermarket availability and secondary outcomes were similar to the results with DIS. The association of our primary relative measure of fast food restaurant availability with Mediterranean Diet score was null, but the association with diabetes status was positive.

DISCUSSION

In this study, we sought to understand the implications of using different geospatial definitions of the food environment on estimating associations with nutrition-related outcomes in epidemiological studies. We found that the choice of buffer-based measure did not change the direction or magnitude of associations with DIS. This is encouraging because our “classic” network buffer was constructed using the administrative unit where a participant resided, which is advantageous for studies without access to exact residential addresses of participants, due to privacy concerns and/or large geographic catchment areas. This finding is also consistent with a recent review study, which reported that different buffer types did not influence effect size and significance in studies of the food environment and obesity.(5)

The review study, however, also reported that different buffer sizes do not influence effect size.(5) We instead found that the strength of associations between the food environment and DIS increased with buffer size. A possible explanation is that food outlet availability for participants residing in suburban and rural areas – who travel farther to access their primary food retailer – is under-counted when using smaller buffer sizes, and over-counted when using larger buffer sizes for participants residing in urban areas.(22) We also found that effect estimates derived from census tract-based measures were smaller than those derived from buffer-based measures tailored to community type, potentially because census tract size is typically smaller than larger buffer sizes. This is important for future research because associations may be under- or over-estimated if researchers do not determine *a priori* which buffer size aligns with participants’ travel behaviors. The difference will depend on the extent of the mismatch in the size of administrative and empirically-derived geographic units, which may vary in studies with participants who reside in different community types. Thus, in future studies, we recommend that researchers tailor buffer size to participants’ community type, and only use administrative units if area size aligns with travel behaviors.

The direction and magnitude of associations using relative measures differed greatly from those using density measures. Absolute measures are popular choices in previous food environment studies, but our findings suggest that the relative and absolute availability of specific food outlets may play different roles in affecting diet practices, potentially due to conceptual differences (e.g., relative measures indicate that eating a certain type of food is common). In addition, we observed a high correlation between our density measures of supermarket and fast food restaurant availability, similar to previous work,(7) which makes it challenging to disentangle and interpret their independent associations with DIS. For example, a previous study found that the density of fast food restaurants was associated with a lower risk of being overweight, but the opposite finding for the proportion of fast food restaurants.(29) This statistical artifact highlights a potential reason for inconsistent results from previous studies using absolute measures.(1)

Substantively, we found that a 10% increase in the percentage of fast food restaurants using a “classic” network buffer was associated with a 6.3 point increase in DIS, or approximately 2.5 standard deviations. This is consistent with previous work showing that living in areas with a relatively high proportion of unhealthy food establishments is associated with less healthy dietary habits and obesity.(3, 30) Although no previous studies have examined the relationship between the food environment and DIS, our findings are consistent with previous studies examining effects on other inflammation-related outcomes, such as inflammation biomarkers and DNA methylation of inflammation-related genes.(31, 32) We observed similar findings with respect to diabetes status, which suggests that our findings are robust to choice of inflammation-related outcome. Etiologically, it is possible that poor diet practices may increase risk of diabetes by increasing levels of systemic inflammation.(14)

This study had several limitations, including that we did not directly address whether the constructs are more valid using primary data sources (e.g., ground-truthing), such as measures based on GPS tracking or activity diaries. The latter more accurately and comprehensively capture individuals' environmental exposures, including residential and non-residential destinations, than proxies, but are also more challenging and costly to collect.(29, 33, 34) There was also a mismatch in the buffer sizes of our empirically-derived measures due to researchers leveraging unique evidence to assign travel distances. In addition, we did not assess accessibility measures (e.g., proximity), which may be less sensitive to differences in buffer size; and we did not assess differences in in-store offerings or multiple definitions of food outlets (e.g., chain vs. non-chain), which previous work has found to influence associations with weight status.(8) Our primary strength, however, included leveraging multiple dimensions of the food environment, including multiple sizes and configurations of geographic units. We also based our choice of exposures and outcomes on *a priori* knowledge of how individuals navigate the food environment and the etiology of diabetes, respectively. Yet, we lacked longitudinal data, so we must be cautious about interpreting our cross-sectional associations.

CONCLUSIONS

The constellation of choices for defining the food environment presents a unique challenge to researchers tasked with characterizing its relationship to diet and health outcomes. In our study, the choice of buffer-based measure did not change the direction or magnitude of associations, which is encouraging for researchers who lack exact residential addresses. Based on our finding that estimates differed by buffer size, researchers should tailor their choice of buffer type to community type, or other proxies for participants' travel behaviors. Our findings also suggest that the mix of restaurant types had a negative association with diet inflammation and diabetes, which supports implementing policies and interventions designed to promote healthy food options, including a combination of minimizing unhealthy restaurants and improving the foods sold within restaurants.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This research was conducted by the Diabetes LEAD Network, funded through CDC cooperative agreements U01DP006293 (Drexel University), U01DP006296 (Geisinger-Johns Hopkins University), U01DP006299 (New

York University School of Medicine), and U01DP006302 (University of Alabama at Birmingham), in collaboration with the U.S. Centers for Disease Control and Prevention, Division of Diabetes Translation.

FUNDING

This research was conducted by the Diabetes LEAD Network, funded through CDC cooperative agreements U01DP006293 (Drexel University), U01DP006299 (New York University School of Medicine), and U01DP006302 (University of Alabama at Birmingham), along with collaboration with the U.S. Centers for Disease Control and Prevention, Division of Diabetes Translation. This work was also supported by cooperative agreement U01 NS041588 co-funded by the National Institute of Neurological Disorders and Stroke (NINDS) and the National Institute on Aging (NIA), National Institutes of Health (NIH), Department of Health and Human Services. Additional funding was provided by the NINDS of NIH under award number 1R01NS092706, the NIA of the NIH under award numbers 1R01AG049970 and 3R01AG049970–04S1, the Commonwealth Universal Research Enhancement (C.U.R.E) program funded by the Pennsylvania Department of Health - 2015 Formula award - SAP #4100072543, the Urban Health Collaborative at Drexel University, and the Built Environment and Health Research Group at Columbia University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NINDS or the NIA. Representatives of the NINDS were involved in the review of the manuscript but were not directly involved in the collection, management, analysis or interpretation of the data. The authors thank the other investigators, the staff, and the participants of the REGARDS study for their valuable contributions. A full list of participating REGARDS investigators and institutions can be found at: https://www.uab.edu/soph/regardsstudy/.

ABBREVIATIONS:

Diabetes LEAD	Location, Environmental Attributes, and Disparities Network
DIS	Dietary Inflammation Score
FFQ	Food Frequency Questionnaire
FoodAPS	National Household Food Acquisition and Purchase Survey
NETS	National Establishment Time Series
NSEE	Neighborhood socioeconomic environment
RECVD	Retail Environment and Cardiovascular Disease
REGARDS	Reasons for Geographic and Racial Differences in Stroke

REFERENCES

1. Caspi CE, Sorensen G, Subramanian SV, et al. The local food environment and diet: a systematic review. Health Place. 2012;18(5):1172–87. [PubMed: 22717379]
2. Malambo P, Kengne AP, De Villiers A, et al. Built Environment, Selected Risk Factors and Major Cardiovascular Disease Outcomes: A Systematic Review. PLoS One. 2016;11(11):e0166846. [PubMed: 27880835]
3. Rummo PE, Guilkey DK, Ng SW, et al. Understanding bias in relationships between the food environment and diet quality: the Coronary Artery Risk Development in Young Adults (CARDIA) study. J Epidemiol Community Health. 2017;71(12):1185–90. [PubMed: 28983065]
4. Houston D. Implications of the modifiable areal unit problem for assessing built environment correlates of moderate and vigorous physical activity. Applied Geography. 2014;50:40–7.
5. Gamba RJ, Schuchter J, Rutt C, et al. Measuring the food environment and its effects on obesity in the United States: a systematic review of methods and results. J Community Health. 2015;40(3):464–75. [PubMed: 25326425]

6. Bivoltsis A, Cervigni E, Trapp G, et al. Food environments and dietary intakes among adults: does the type of spatial exposure measurement matter? A systematic review. Int J Health Geogr. 2018;17(1):19. [PubMed: 29885662]
7. Clary CM, Ramos Y, Shareck M, et al. Should we use absolute or relative measures when assessing foodscape exposure in relation to fruit and vegetable intake? Evidence from a wide-scale Canadian study. Prev Med. 2015;71:83–7. [PubMed: 25481095]
8. Wilkins E, Morris M, Radley D, et al. Methods of measuring associations between the Retail Food Environment and weight status: Importance of classifications and metrics. SSM Popul Health. 2019;8:100404. [PubMed: 31245526]
9. Howard VJ, Cushman M, Pulley L, et al. The reasons for geographic and racial differences in stroke study: objectives and design. Neuroepidemiology. 2005;25(3):135–43. [PubMed: 15990444]
10. Block G, Woods M, Potosky A, et al. Validation of a self-administered diet history questionnaire using multiple diet records. J Clin Epidemiol. 1990;43(12):1327–35. [PubMed: 2254769]
11. Boucher B, Cotterchio M, Kreiger N, et al. Validity and reliability of the Block98 food-frequency questionnaire in a sample of Canadian women. Public Health Nutr. 2006;9(1):84–93. [PubMed: 16480538]
12. Hirsch AG, Carson AP, Lee NL, et al. The Diabetes Location, Environmental Attributes, and Disparities Network: Protocol for Nested Case Control and Cohort Studies, Rationale, and Baseline Characteristics. JMIR Res Protoc. 2020;9(10):e21377. [PubMed: 33074163]
13. Byrd DA, Judd SE, Flanders WD, et al. Development and Validation of Novel Dietary and Lifestyle Inflammation Scores. J Nutr. 2019;149(12):2206–18. [PubMed: 31373368]
14. Calle MC, Fernandez ML. Inflammation and type 2 diabetes. Diabetes & Metabolism. 2012;38(3):183–91. [PubMed: 22252015]
15. Tsivgoulis G, Judd S, Letter AJ, et al. Adherence to a Mediterranean diet and risk of incident cognitive impairment. Neurology. 2013;80(18):1684–92. [PubMed: 23628929]
16. Messer LC, Laraia BA, Kaufman JS, et al. The development of a standardized neighborhood deprivation index. J Urban Health. 2006;83(6):1041–62. [PubMed: 17031568]
17 . McAlexander TP, Algur Y, Schwartz BS, Rummo PE, Lee D, Siegel KR, Ryan V, Lee NL, Malla G, McClure LA Categorizing community type for epidemiologic evaluation of community factors and chronic disease across the United States 2021 Unpublished.
18. Drexel University Urban Health Collaborative. The Retail Environment and Cardiovascular Disease (RECVD) Project. https://sites.google.com/view/recvd-team-project-site.
19. Hirsch JA, Moore KA, Cahill J, et al. Business Data Categorization and Refinement for Application in Longitudinal Neighborhood Health Research: a Methodology. Journal of Urban Health. 2020.
20. U.S. Census Bureau 2016 TIGER Line Geodatabase files.
21. Forsyth A: LEAN GIS (Local Environment for Activity and Nutrition) Protocols.Version 2.1. 2012, http://designforhealth.net/wp-content/uploads/2012/12/LEAN_Protocol_V2_1_010112rev.pdf.
22. National Household Food Acquisition and Purchase Survey (FoodAPS). Economic Research Service (ERS), U.S. Department of Agriculture (USDA). https://www.ers.usda.gov/foodaps. Accessed April 24, 2019.
23. Berke EM, Koepsell TD, Moudon AV, et al. Association of the built environment with physical activity and obesity in older persons. Am J Public Health. 2007;97(3):486–92. [PubMed: 17267713]
24. Duncan M, Mummery K. Psychosocial and environmental factors associated with physical activity among city dwellers in regional Queensland. Prev Med. 2005;40(4):363–72. [PubMed: 15530589]
25. Frank L, Kerr J, Chapman J, et al. Urban form relationships with walk trip frequency and distance among youth. Am J Health Promot. 2007;21(4 Suppl):305–11. [PubMed: 17465175]
26. Moudon AV, Lee C, Cheadle AD, et al. Operational Definitions of Walkable Neighborhood: Theoretical and Empirical Insights. J Phys Act Health. 2006;3(s1):S99–s117. [PubMed: 28834523]
27. Sallis JF, Hovell MF, Hofstetter CR, et al. Distance between homes and exercise facilities related to frequency of exercise among San Diego residents. Public Health Rep. 1990;105(2):179–85. [PubMed: 2108465]

28. Thornton LE, Lamb KE, White SR. The use and misuse of ratio and proportion exposure measures in food environment research. Int J Behav Nutr Phys Act. 2020;17(1):118. [PubMed: 32957988]
29. Kestens Y, Lebel A, Chaix B, et al. Association between activity space exposure to food establishments and individual risk of overweight. PLoS One. 2012;7(8):e41418. [PubMed: 22936974]
30. Cooksey-Stowers K, Schwartz MB, Brownell KD. Food Swamps Predict Obesity Rates Better Than Food Deserts in the United States. Int J Environ Res Public Health. 2017;14(11).
31. Giurgescu C, Nowak AL, Gillespie S, et al. Neighborhood Environment and DNA Methylation: Implications for Cardiovascular Disease Risk. J Urban Health. 2019;96(Suppl 1):23–34. [PubMed: 30635842]
32. King K. Neighborhood walkable urban form and C-reactive protein. Prev Med. 2013;57(6):850–4. [PubMed: 24096140]
33. Sadler RC, Gilliland JA. Comparing children’s GPS tracks with geospatial proxies for exposure to junk food. Spatial Spatio-temporal Epidemiology. 2015;14:55–61. [PubMed: 26530823]
34. Perchoux C, Chaix B, Brondeel R, et al. Residential buffer, perceived neighborhood, and individual activity space: New refinements in the definition of exposure areas - The RECORD Cohort Study. Health Place. 2016;40:116–22. [PubMed: 27261634]
35. Forsyth A, Van Riper D, Larson N, et al. Creating a replicable, valid cross-platform buffering technique: The sausage network buffer for measuring food and physical activity built environments. International Journal of Health Geographics. 2012;11(1):14. [PubMed: 22554353]

HIGHLIGHTS

- Relative availability of fast food restaurants was positively associated with DIS.
- Relative availability of supermarkets, however, was not associated with DIS.
- The difference in associations across person-based buffers was negligible.
- Researchers should tailor buffer-based measures to community type in future work.

TABLE 1.

Food environment measures: data sources, definitions, and coverage

	Census tract	"Classic" network buffer	"Sausage" network buffer
Neighborhood definition	Administrative	Administrative	Person-based[1]
Input point	N/A	Population-weighted centroid of census tract	Residential address
Size of geographic unit	Fixed	1-mile (1.6 km) buffer[2] 2-mile (3.2 km) buffer 6-mile (10 km) buffer 10-mile (16 km) buffer	1-km buffer 5-km buffer 8-km buffer 16-km buffer
Geographic measures	Density (count per square kilometer) Percentage (count per total food stores or restaurants)	Density (count per square kilometer) Percentage (count per total food stores or restaurants)	Density (count per square kilometer) Percentage (count per total food stores or restaurants)
Geographic coverage	Contiguous U.S. (72,538 (99.3%[3]) CTs)	Contiguous US (72,538 (99.3%[3]) CTs)	REGARDS participants' home addresses (13,914 (19%[3]) CTs)
Geoprocessing – polygon construction	Columbia University[4]	New York University[5]	University of Michigan[6]
Geoprocessing – spatial join	Columbia University[4]	Drexel University[5]	Columbia University[4]

NOTE: All measures use NETS establishment data to define the supermarket and fast food restaurant exposure measures.

[1] A person-based neighborhood definition refers to a neighborhood as centered around a person-specific location (e.g., home) and across a distance that aims to capture the local environment as experienced or imagined by a participant.(35) The University of Michigan geocoded REGARDS participants to their residential addresses.

[2] 1-mile "classic" network buffer was constructed using the 'walking' travel mode; the remaining network buffers used the 'driving' travel mode.

[3] Denominator is continental U.S., which includes the 50 states and the District of Columbia.

[4] The Built Environment and Health (BEH) Working Group at Columbia University buffered REGARDS participants' residential addresses by 1- and 5-km (straight-line distance) and geoprocessed NETS establishment data within census tract and "sausage" network buffer boundaries. Before geoprocessing, areas of hydrography were removed to accurately calculate land area. Aggregate counts for "sausage" network buffers exclude establishments with low geocoding quality (i.e. street name, zip code, city centroid ~ 8%).

[5] The Department of Population Health at New York University Grossman School of Medicine buffered population-weighted centroids of census tracts where REGARDS participants lived at baseline by 1-, 2-, 6-, and 10-mile (street network distance). The "classic" network buffer was created using ESRI's ArcGIS StreetMap Premium 2019 for the street network data and the "generalized" polygon option and default settings (with default trim and standard precision) in ArcGIS Pro 2.4.2. The Diabetes LEAD Network Data Coordinating Center at Data Drexel University geoprocessed NETS establishment data within "classic" network buffer.

[6] The University of Michigan buffered REGARDS participants' residential addresses by 1-, 5-, 8-, and 16-km (street network distance using a 150-meter radius from the street centerline. The "sausage" network buffer was created using ESRI's ArcGIS StreetMap Premium 2017 (North America HERE Release 3) for the street network data and following these buffer specifications in ArcGIS Pro 2.1 or later.(35)

TABLE 2.

Descriptive statistics of the REGARDS sample

	N or median or mean	% or IQR or SD
Gender[a]		
Male	11,353	55.8%
Female	8,978	44.2%
Education[a]		
Less than high school	1,923	9.5%
High school graduate	5,159	25.4%
Some college	5,571	27.4%
College graduate and above	7,670	37.7%
Race[a]		
Black	6,824	33.6%
White	13,507	66.4%
Income[a]		
<$20,000	4,910	24.2%
$20,000-$34,000	6,409	31.5%
$35,000-$74,000	3,533	17.4%
$75,000 and above	2,318	11.4%
Refused	3,161	15.5%
NSEE[a]	18.7	10.9, 28.8
Community type		
Higher-density urban	3,389	16.7%
Lower-density urban	8,093	39.8%
Suburban/small town	3,921	19.3%
Rural	4,928	24.2%
DIS[b]	−0.004	2.52
Mediterranean Diet Score[c]	4.4	1.7
Diabetes status		
No	22,430	75.0%
Yes	6,345	21.2%
Missing	1,121	3.7%

NOTE: IQR=interquartile range; SD=standard deviation

[a] Sample excludes those with missing DIS.

[b] Theoretical range: 0 –9.

[c] Theoretical range: −14.9–12.8.

TABLE 3.

Descriptive statistics of food environment measures, by geographic definition, food outlet type, and buffer size (n=20,331)

	Census tract	**“Classic” network buffer**	**Sausage network buffer**
	Mean (SD)	**Mean (SD)**	**Mean (SD)**
Supermarkets			
Density, tailored[a]	0.12 (0.48)	0.22 (0.40)	0.22 (0.46)
Density, 1 km[b]		0.26 (0.45)	0.25 (0.57)
Density, 5 km[c]		0.25 (0.36)	0.22 (0.34)
Density, 8 km		0.18 (0.28)	0.20 (0.31)
Density, 16 km		0.14 (0.23)	0.16 (0.25)
Percentage, tailored[a]	6.9 (11.7)	11.7 (7.8)	11.6 (8.2)
Percentage, 1 km[b]		8.1 (13.0)	6.9 (15.7)
Percentage, 5 km[c]		10.4 (11.0)	10.9 (8.9)
Percentage, 8 km		11.2 (6.3)	11.1 (7.0)
Percentage, 16 km		11.2 (4.4)	11.2 (4.4)
Fast food restaurants			
Density, tailored[a]	0.55 (1.19)	0.69 (0.90)	0.66 (0.95)
Density, 1 km[b]		0.74 (1.10)	0.67 (1.35)
Density, 5 km[c]		0.77 (0.86)	0.71 (0.74)
Density, 8 km		0.59 (0.63)	0.65 (0.66)
Density, 16 km		0.48 (0.56)	0.55 (0.58)
Percentage, tailored[a]	27.4 (23.9)	32.6 (14.6)	32.0 (15.2)
Percentage, 1 km[b]		21.3 (23.4)	16.0 (25.0)
Percentage, 5 km[c]		27.7 (19.1)	29.7 (16.3)
Percentage, 8 km		31.2 (12.2)	30.9 (13.4)
Percentage, 16 km		31.6 (9.3)	31.5 (9.5)

NOTE: 1-km and 8-km network buffer measures were calculated using walking distance, and other distances were calculated using driving distance.

[a] Measure is tailored to the community type (high-density urban, low-density urban, suburban/small town, rural) of participants’ residential census tract.

[b] “Classic” network buffer distances for “one-size-fits-all” measures are 1 mile (~1.6 km), 2 miles (~3.2 km), 6 miles (~9.7 km), and 10 miles (~16.1 km).

TABLE 4.

Model-based associations between food environment measures and diet inflammation score, by geographic definition, food outlet type, and buffer size (n=20,331)

	Census tract		"Classic" network buffer[a]		Sausage network buffer	
	β (SE)	p-value	β (SE)	p-value	β (SE)	p-value
Supermarkets						
Density, tailored[b]	**−0.12 (0.03)**	**0.001**	**−0.30 (0.05)**	**<0.001**	**−0.25 (0.04)**	**<0.001**
Density, 1 km			**−0.19 (0.04)**	**<0.001**	**−0.12 (0.03)**	**<0.001**
Density, 5 km			**−0.27 (0.06)**	**<0.001**	**−0.34 (0.07)**	**<0.001**
Density, 8 km			**−0.46 (0.08)**	**<0.001**	**−0.39 (0.07)**	**<0.001**
Density, 16 km			**−0.58 (0.08)**	**<0.001**	**−0.53 (0.08)**	**<0.001**
Percentage, tailored[b]	−0.01 (0.13)	0.97	0.42 (0.22)	0.05	0.29 (0.20)	0.16
Percentage, 1 km			0.05 (0.12)	0.65	0.04 (0.10)	0.68
Percentage, 5 km			0.26 (0.14)	0.06	0.20 (0.18)	0.26
Percentage, 8 km			0.38 (0.26)	0.14	0.39 (0.22)	0.07
Percentage, 16 km			0.63 (0.36)	0.08	**1.03 (0.37)**	**0.01**
Fast food restaurants						
Density, tailored[b]	**−0.08 (0.02)**	**<0.001**	**−0.15 (0.02)**	**<0.001**	**−0.14 (0.02)**	**<0.001**
Density, 1 km			**−0.09 (0.02)**	**<0.001**	**−0.05 (0.01)**	**<0.001**
Density, 5 km			**−0.16 (0.02)**	**<0.001**	**−0.21 (0.03)**	**<0.001**
Density, 8 km			**−0.26 (0.03)**	**<0.001**	**−0.25 (0.03)**	**<0.001**
Density, 16 km			**−0.30 (0.03)**	**<0.001**	**−0.29 (0.03)**	**<0.001**
Percentage, tailored[b]	**0.27 (0.06)**	**<0.001**	**0.63 (0.12)**	**<0.001**	**0.52 (0.11)**	**<0.001**
Percentage, 1 km			0.12 (0.07)	0.07	**0.18 (0.06)**	**0.004**
Percentage, 5 km			**0.26 (0.08)**	**0.002**	**0.42 (0.10)**	**<0.001**
Percentage, 8 km			**0.5 (0.13)**	**<0.001**	**0.53 (0.12)**	**<0.001**
Percentage, 16 km			**1.02 (0.19)**	**<0.001**	**0.81 (0.18)**	**<0.001**

NOTE: Bold denotes statistically significant at α <0.05 level. Supermarkets and fast food restaurants were included in separate regression models; and we controlled for total food outlets (continuous) in models with relative measures.

[a] Measure is tailored to the community type (high-density urban, low-density urban, suburban/small town, rural) of participants' residential census tract.

[b] "Classic" network buffer distances for "one-size-fits-all" measures are 1 mile (~1.6 km), 2 miles (~3.2 km), 6 miles (~9.7 km), and 10 miles (~16.1 km).