BMJ Open

Geographic scale matters in detecting the relationship between neighbourhood food environments and obesity risk: an analysis of driver license records in Salt Lake County, Utah

Jessie X Fan,[1] Heidi A Hanson,[2] Cathleen D Zick,[1] Barbara B Brown,[1] Lori Kowaleski-Jones,[1] Ken R Smith[3]

CrossMark

For numbered affiliations see end of article.

Correspondence to
Dr Jessie X Fan;
fan@fcs.utah.edu

ABSTRACT

Objectives: Empirical studies of the association between neighbourhood food environments and individual obesity risk have found mixed results. One possible cause of these mixed findings is the variation in neighbourhood geographic scale used. The purpose of this paper was to examine how various neighbourhood geographic scales affected the estimated relationship between food environments and obesity risk.

Design: Cross-sectional secondary data analysis.

Setting: Salt Lake County, Utah, USA.

Participants: 403 305 Salt Lake County adults 25–64 in the Utah driver license database between 1995 and 2008.

Analysis: Utah driver license data were geo-linked to 2000 US Census data and Dun & Bradstreet business data. Food outlets were classified into the categories of large grocery stores, convenience stores, limited-service restaurants and full-service restaurants, and measured at four neighbourhood geographic scales: Census block group, Census tract, ZIP code and a 1 km buffer around the resident's house. These measures were regressed on individual obesity status using multilevel random intercept regressions.

Outcome: Obesity.

Results: Food environment was important for obesity but the scale of the relevant neighbourhood differs for different type of outlets: large grocery stores were not significant at all four geographic scales, limited-service restaurants at the medium-to-large scale (Census tract or larger) and convenience stores and full-service restaurants at the smallest scale (Census tract or smaller).

Conclusions: The choice of neighbourhood geographic scale can affect the estimated significance of the association between neighbourhood food environments and individual obesity risk. However, variations in geographic scale alone do not explain the mixed findings in the literature. If researchers are constrained to use one geographic scale with multiple categories of food outlets, using Census tract or 1 km buffer as the neighbourhood

---

Strengths and limitations of this study

- Our study applies four neighbourhood geographic scales utilising the Utah driver license data, which allows us to detect the role of geographic scale in the relationship between obesity and neighbourhood food environment.
- Our study provides evidence that the significance of the association between obesity and neighbourhood food environment depends on the neighbourhood geographic scale used and the specific element of the food environment that is in question. If the relevant neighbourhood is defined too large (ie, larger than Census tract for convenience stores and full-service restaurants) or too small (ie, smaller than Census tract for limited-service restaurants), then the statistical relationship is insignificant.
- The most important limitation of our study is the cross-sectional nature of our data, which prevents us from inferring causal relationships between elements of the neighbourhood food environment and obesity.
- Another important limitation is that we used self-reported weight and height data, which could systematically underestimate true obesity status. However, as long as such self-reporting bias is not geographically patterned, the multivariate results should not be affected.

geographic unit is likely to allow researchers to detect most significant relationships.

BACKGROUND

The growing obesity epidemic in the USA[1] [2] has motivated policymakers and researchers

to search for modifiable factors that may be linked to individual weight. One of the modifiable factors is the neighbourhood food environment. The neighbourhood food environment involves the "number, type, location, and accessibility of food outlets in a location,"[3] with food outlets predominantly including stores (supermarkets, grocery stores and convenience stores) and restaurants (full-service and limited-service, including fast food). Early research suggested that diet quality can be improved, and unhealthy weight gain can be prevented through better access to supermarkets and large grocery stores and reduced exposure to limited-service restaurants and convenience stores.[4–6] As a result, public policy efforts to reduce obesity have targeted the food environment as an important area for intervention. For example, in 2008 the City of Los Angeles passed a bill to ban the opening of any fast food restaurants in the poor neighbourhoods in the city.[7] The White House Childhood Obesity Task Force proposed that communities increase the number of supermarkets in order to reduce childhood obesity.[8] However, as more research has become available based on different study locations using diverse data sources, measurements and analytical methods, the relationship between neighbourhood food environment and obesity has been challenged. While some studies have supported earlier findings,[9 10] others have found null or even opposite relationships.[11 12] Inconsistencies in study results are not surprising given the variations in neighbourhood food environment measures, study locations, data sources, analytical methods and controls for confounding factors.

One important potential contributor to mixed findings is the geographic scale of the neighbourhood used in studies. Geographic scale can influence the observed relationship between neighbourhood context and individuals' health outcomes,[13] which is an example of the well-known modifiable areal unit problem (MAUP) associated with the use of geographically aggregated data.[14 15] Finding appropriate and consistent criteria for defining geographic boundaries of a neighbourhood can be challenging because the appropriate definitions of neighbourhood probably differ across places, environmental factors and outcomes being studied. For example, large grocery stores are likely to draw patrons from a larger geographic area than convenience stores so the aggregation level appropriate for large grocery stores might not be suitable for convenience stores.

The scale of US neighbourhood in the food environment literature has varied considerably, from Census block group,[16 17] Census tract,[18 19] ZIP code[6] to state.[20] When a buffer method was used, there was a wide range of buffer distances, ranging from 100 m[21] to 5 km,[22] with two studies examining multiple distances for what is considered walkable (eg, 500 m and 1 km) to needing transportation (eg, 2 and 5 km).[22 23] Several recent studies also used store density within a defined geographic area per 100 000 people as the geographic measure of interest.[24 25]

In this study we examined how four geographic scales of neighbourhood food environment might affect the direction, size and significance of the relationship between neighbourhood retail food establishments and individual obesity risk using population and business data in Salt Lake County, Utah. Three of these four were area measures: Census block group, Census tract and ZIP code; and the fourth was a buffer measure, which was a 1 km buffer around an individual's residence. The 1 km buffer was chosen because of its proven usefulness in past build environment-related research and its compatibility with our intersection density measure.[22 26] Because the relevant geographic scale is likely to vary by outlet type, we included four types of food retail outlets: large grocery stores, convenience stores, limited-service restaurants and full-service restaurants. We related these various measures of the neighbourhood food environment to individual obesity risk while controlling for other individual and neighbourhood factors in an attempt to understand the role of geographic scale in the estimated relationship between neighbourhood food environment and individual obesity risk.

METHODS

Our study area was Salt Lake County, the county with the largest population in the state of Utah. Individual height and weight and demographic data were taken from the Utah Population Database (UPDB),[27] which contains driver license data from the Utah Driver License Division (DLD) going back to 1985. Individual DLD information is updated at the time of a license renewal every 5 years and when a person moves. The UPDB staff linked driver license data (residential location, height, weight, gender and age) to other data via Universal Transverse Mercator (UTM) coordinates. All other personal information from the DLD was removed before the data were provided to the investigators to protect the confidentiality of the individuals.

We excluded young adults (younger than 25) who had likely not established their postadolescence residence. We also excluded elderly adults (older than 64) whose body mass indices (BMIs) were likely to have more complex associations with health conditions.[28] In addition, we excluded residents who lived in sparsely settled fringe areas of Salt Lake County where the physical environment was mountainous and very different from the urban setting. Because the DLD data had multiple records for many drivers, we selected the DLD record closest to year 2000 to synchronise with the 2000 Census for any given individual. From the driver license data, self-reported height and weight information were converted to BMI (which is (weight in kg)/(height in m)^2), with obesity defined as having a BMI $\geq$30. BMI data obtained from driver licenses have the advantage of extensive coverage but the potential limitation of self-report bias, although there was no reason to suspect that such self-report bias would be geographically patterned.

With these restrictions in place, the resulting sample size comprised of 403 305 individuals. According to the US Census, Salt Lake County had 436 438 persons aged 25–64 in 2000 so our sample had about 92.4% of the age-eligible Salt Lake County population.[29]

UPDB data were linked to the 2000 US Census[30] data that measured neighbourhood demographic, socio-economic and walkability characteristics at the Census block group, Census tract and ZIP code levels. Demographic and socioeconomic control variables included median neighbourhood income, distinct percentages of blacks, Hispanics, Asians and Hawaiians/Pacific Islanders, and median age of residents. Walkability measures included population density, intersection density derived from a 1 km network buffer from the US Census TIGER/Line file,[31] and median age of neighbourhood housing because older neighbourhoods have typically been designed with more walkable features such as tree-shaded sidewalks, narrow streets and mixed land use.[17] [32–34] Combined, these latter measures captured neighbourhood population density, land use diversity and pedestrian-friendly design, the '3Ds' of walkability.[35]

Data on neighbourhood food environments came from the 2000 Dun & Bradstreet files[36] and were linked to DLD and Census data through geocoding. Businesses with food-related Standard Industrial Classification (SIC) codes were selected and grouped into four categories: (1) large grocery stores, including general grocery stores, supermarkets, independent and chain grocery stores and fish and vegetable markets, (2) convenience stores, including both chain and independent convenience stores, and drug stores, (3) limited-service food restaurants, including drive-in restaurants, chain and independent fast food restaurants, and pizza restaurants and (4) full-service restaurants, including various American and ethnic restaurants, grills, cafeteria, seafood, steak, barbeque restaurants and buffets. Businesses within each category were aggregated for four geographic units: (1) Census block group, (2) Census tract, (3) ZIP code and (4) a 1 km buffer measure around each individual's home. In total, there were 550 Census block groups, 190 Census tracts and 32 ZIP codes included in our sample. Figure 1 illustrates an example of the relative size of Census block groups, Census tracts, ZIP codes and a 1 km buffer in Salt Lake County, Utah.

Multilevel regression models were estimated to investigate how associations between neighbourhood food environments and individual obesity risk varied across different geographic scales of the neighbourhood. All four retail outlet types were included in each model at each scale. Random intercept models with two levels were estimated: person level and neighbourhood level, with the neighbourhood level being Census block group, tract and ZIP code, respectively. For the 1 km buffer model, Census block group was defined as the neighbourhood for the control variables because the

average area size for Census block groups was the closest to a 1 km buffer area. Control variables based on past studies[17] [34] included individual age, median household income and the percentages of African-Americans, Hispanics, Asians and Hawaiians/Pacific Islanders measured at the relevant neighbourhood scale. Walkability measures included intersection density for a 1 km buffer around residents' addresses, population density and median housing age at block group, tract and ZIP code levels. Each model was estimated for men and women separately.[34] [37] Analyses were conducted in SAS V.9.2 (Proc Glimmix). Collinearity diagnostics revealed no problematic multicollinearity among the independent variables. Using the tract-level model as an example, the highest variance inflation factor (VIF) was 2.57 for per cent of Hispanics in the neighbourhoods, and the highest condition index was 3.58, well within the acceptable range for multicollinearity diagnosis. For the four food outlet variables, the VIF ranged from 1.17 for healthy grocery stores to 2.57 for full-service restaurants.

RESULTS

Descriptive statistics for individual obesity status and control variables are shown in table 1. The estimated obesity rate was 17.16% for adults 25–64 in Salt Lake County around year 2000, with 18.28% for men and 15.94% for women. This is in line with obesity rate estimates of between 15% and 19% for Utah from the 2001 Behavioral Risk Factor Surveillance System (BRFSS).[38] Neighbourhood statistics at the Census tract level are presented while statistics at other geographic scales are available from the authors on request. Table 2 presents average sizes of each definition of the neighbourhood and population-weighted mean counts and SDs of retail food outlets. On average, Census block group was the smallest geographic unit with the least number of food outlets in all four categories, followed by Census tract, 1 km buffer and ZIP code. Regardless of the geographic scale, the number of full-service restaurants was the highest among all four food outlet categories, followed by limited-service restaurants and convenience stores. For every grocery store, there were approximately 7.5 full-service restaurants, 2.5 limited-service restaurants and 2 convenience stores. Correlation coefficients between food environment measures at different geographic scales ranged from a low of 0.15 for convenience stores measured at the Census tract and ZIP code levels to a high of 0.64 for full-service restaurants measured at the Census block group and Census tract levels (not shown in tables but available from the authors on request).

Table 3 presents the estimated OR, 95% confidence limits, and p values for the various measures of the food environment as they relate to individual obesity risk, controlling for the list of covariates discussed earlier. On the one hand, if Salt Lake County was considered the study population, then our study covered almost all the

Figure 1 An example of the relative size of Census block groups, tracts, ZIP codes and 1 km buffer in Salt Lake County, Utah.
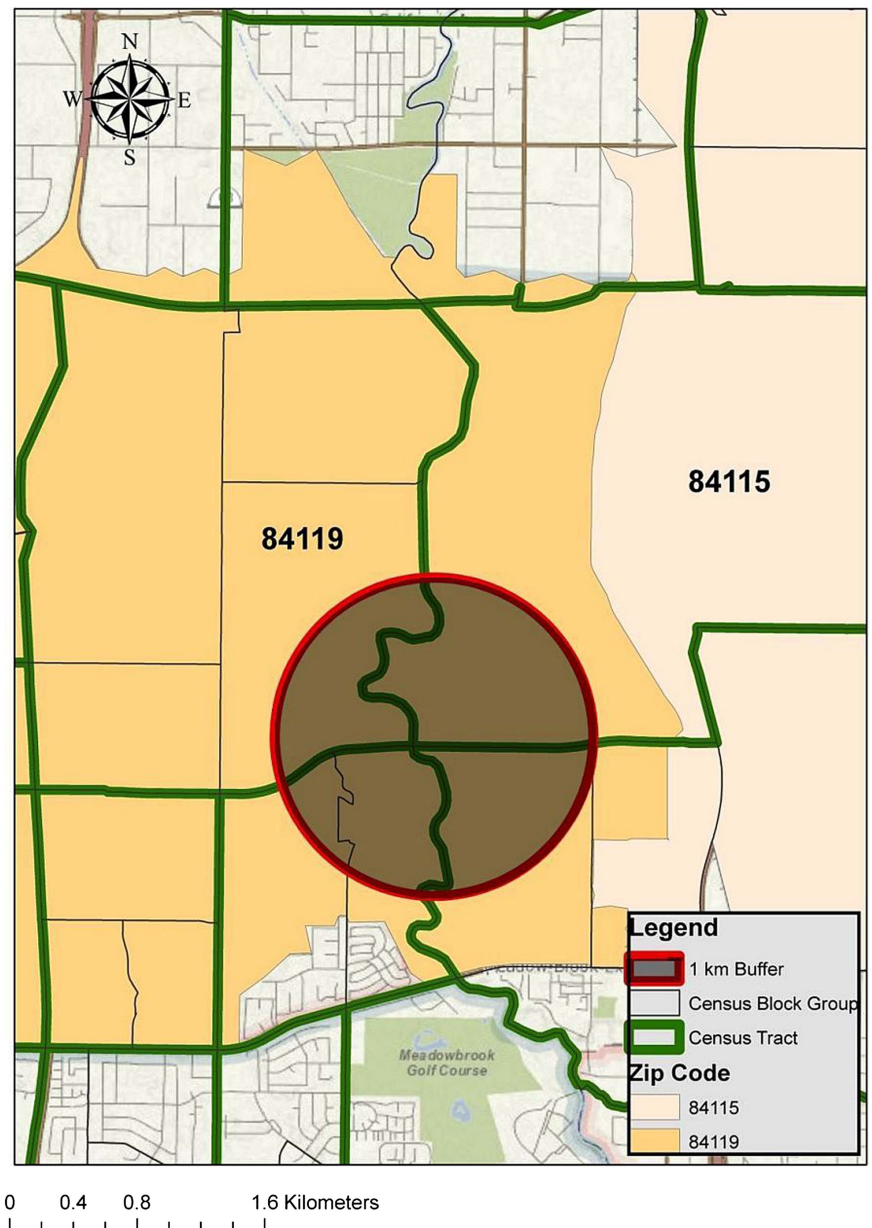
population, making sampling testing statistics such as p values irrelevant. On the other hand, if Salt Lake County was considered a sample representing a larger population (eg, US population residing in similar-sized urban areas), then with such a large sample size, p values can quickly approach zero and as such have a

Table 1 Descriptive statistics for the sample: individual and tract-level statistics

Variable	Whole sample (n=403 305)	Men (n=209 946)	Women (n=193 359)
Obese (%) (individual level)	17.16	18.28	15.94
Mean individual age (SD)	40.52 (10.50)	40.33 (10.44)	40.74 (10.56)
Tract median family income in $1000 (SD)	51.66 (17.22)	51.13 (17.40)	52.23 (17.01)
Tract (%) African-American	0.86	0.89	0.82
Tract (%) Pacific Islander/Hawaiian	1.13	1.16	1.10
Tract (%) Hispanic	11.80	12.21	11.36
Tract (%) Asian	2.62	2.65	2.58
Tract population density/miles^2 in 1000 (SD)	5.26 (2.53)	5.28 (2.56)	5.23 (2.50)
Intersection density 1 km buffer (SD)	41.53 (10.91)	41.46 (10.93)	41.62 (10.88)
Tract median housing age in years (SD)	26.73 (14.84)	27.03 (14.96)	26.41 (14.70)

Table 2 Counts of food outlets by geographic scale: means and SD

Food store type	Block group	Tract	ZIP code	1 km buffer
Average area size in km² (SD)	1.24 (2.32)	8.51 (32.68)	381.84 (781.15)	3.14 (no SD)
Large grocery stores (SD)	0.16 (0.46)	0.46 (0.73)	2.83 (1.96)	0.64 (1.02)
Convenience stores (SD)	0.32 (0.66)	1.02 (1.23)	5.95 (3.25)	1.33 (1.72)
Limited-service restaurants (SD)	0.38 (0.90)	1.19 (1.82)	7.74 (5.43)	1.59 (2.28)
Full-service restaurants (SD)	1.19 (2.70)	3.53 (4.67)	21.67 (12.35)	4.79 (6.70)

tendency to reject the null.[39] With this understanding, we note effect sizes, 95% confidence limits and p values as suggested by Lin et al.[39]

Large grocery stores: Table 3 shows that for large grocery stores, the point estimates for the ORs ranged from 0.999 (1 km buffer for men) to 0.966 (block group for women), indicating a beneficial role of large grocery stores in reducing individual obesity risk. For the three area measures, 95% CI narrowed as the size of area increased, with ZIP code level having the narrowest CI and the lowest p value for both genders. However, none of the OR had a p value less than 0.05, indicating statistical insignificance of these associations using conventional sampling statistics.

Convenience stores: For convenience stores, the point estimates for the ORs ranged from 0.986 (ZIP code for women) to 0.957 (block group for men), indicating that larger neighbourhood counts of convenience stores were associated with a lower individual obesity risk. For the three area measures, the effect size decreased but the 95% CI also narrowed as the geographic area size increased. This protective effect existed for both men and women when neighbourhood geographic scale was the Census tract or 1 km buffer. In addition, the association was significant at the Census block group level for men.

Limited-service restaurants: For limited-service restaurants, the point estimates for the ORs ranged from 1.015 (ZIP code for men) to 1.032 (block group for women), indicating that greater neighbourhood counts of limited-service restaurants were associated with a higher individual obesity risk. For the three area measures, the effect size was the largest at the tract level for men and at the block group level for women. The 95% CI narrowed as the size of the area increased. This detrimental relationship was statistically significant at the Census tract, ZIP code and 1 km buffer levels, but not the Census block group level.

Full-service restaurants: For full-service restaurants, the point estimates for the ORs ranged from 0.998 (ZIP code for men) to 0.984 (tract for women), indicating that greater neighbourhood counts of full-service restaurants were associated with a lower individual risk of obesity. For the three area measures, the larger the area size, the smaller the effect for men. However, for women, the effect was the largest at the tract level. Regardless of the effect size, the 95% CI narrowed as area size increased. This protective relationship was statistically significant for both

genders at the Census tract or 1 km buffer levels, but not the ZIP code level. In addition, the relationship was statistically significant for men at the block group level.

For the control variables, the multilevel logistic regression coefficients were in general agreement across different models so only the tract-level results are presented in table 4. Estimates at other geographic scales are available from the authors on request. Older individual age was associated with an increased obesity risk. The neighbourhood socioeconomic and demographic variables were also consistently associated with individual obesity risk. Higher median neighbourhood income was associated with a lower individual obesity risk. Having higher percentages of Hawaii-Pacific islanders and Hispanics in the neighbourhood were associated with a higher individual risk of obesity while having higher percentages of African-Americans (men only) and Asians in the neighbourhood were associated with a lower individual risk of obesity. Because we did not have individual race/ethnicity data, it was likely that at least some of these neighbourhood racial/ethnic effects were artefacts of individual racial/ethnic effects. For neighbourhood physical environment variables, a higher population density and older housing stock in the neighbourhood were both associated with a lower individual risk of obesity, while a higher intersection density in the neighbourhood was associated with a higher individual risk of obesity.

DISCUSSION

The purpose of this study was to investigate the role of geographic scale on the relationship between neighbourhood food environments and individual obesity risk. Our results show that the relationship between obesity risk and different types of food outlets was significant at different levels of geographic scale. Large grocery stores were not statistically significantly associated with individual risk of obesity at all four geographic scales, although the protective effect had a p value of 0.081 for men at the ZIP code level. It is possible that a geographic scale larger than the ZIP code is needed to detect significant relationships between large grocery stores and obesity risk as people often travel substantial distances when doing major grocery shopping. Limited-service restaurants were detrimental at medium levels of geographic scale of our analysis (Census tract or larger, or 1 km buffer). Convenience stores and full-service restaurants

Table 3 OR estimates of multilevel logistic regression analyses for individual obesity risk*

	Block group OR (95% CL)	p Value	Tract OR (95% CL)	p Value	ZIP code OR (95% CL)	p Value	1 km buffer OR (95% CL)	p Value
Men (n=209 946)								
Large grocery	0.966 (0.918, 1.016)	0.1785	0.972 (0.935, 1.011)	0.1626	0.973 (0.944, 1.003)	0.081	0.999 (0.980, 1.018)	0.9189
Convenience	0.957 (0.925, 0.991)	0.0143	0.966 (0.942, 0.991)	0.0088	0.985 (0.970, 1.000)	0.0556	0.959 (0.947, 0.972)	<.0001
Limited-service	1.019 (0.991, 1.048)	0.1748	1.022 (1.002, 1.041)	0.0288	1.015 (1.003, 1.029)	0.0183	1.018 (1.008, 1.028)	0.0004
Full-service	0.987 (0.977, 0.997)	0.0144	0.989 (0.981, 0.998)	0.0131	0.998 (0.993, 1.004)	0.5676	0.993 (0.990, 0.997)	0.0004
Women (n=193 359)								
Large grocery	0.960 (0.896, 1.028)	0.2406	0.970 (0.918, 1.026)	0.2867	0.972 (0.932, 1.014)	0.1866	0.999 (0.976, 1.023)	0.9507
Convenience	0.962 (0.919, 1.007)	0.0982	0.963 (0.928, 0.998)	0.0382	0.986 (0.965, 1.008)	0.2041	0.979 (0.963, 0.994)	0.0082
Limited-service	1.032 (0.995, 1.071)	0.0947	1.031 (1.003, 1.060)	0.0270	1.018 (1.001, 1.036)	0.0438	1.018 (1.006, 1.031)	0.0038
Full-service	0.987 (0.973, 1.000)	0.0567	0.984 (0.972, 0.996)	0.0087	0.994 (0.986, 1.002)	0.1291	0.992 (0.987, 0.997)	0.0007

*Regressions control for individual age, median family income, percentages of residents who were African-American, Pacific Islander/Hawaiian, Hispanic and Asian, population density, street intersection density and median housing age.

were protective at the smallest geographic scale (Census tract or smaller, or 1 km buffer). Table 5 summarises these findings for the four types of food outlets by gender. Although geographic scale affected the significance level and the size of the coefficients, the directions of the coefficients were remarkably consistent across geographic scales and across gender. More counts of large grocery stores, convenience stores and full-service restaurants were protective of individual obesity risk, while more counts of limited-service restaurants were detrimental to individual obesity risk. We also observed that for the three area measures, with the exception of limited-service restaurants, the estimated size of the effect decreased when the neighbourhood geographic scale increased, regardless of the statistical significance of the estimate.

Our findings provide evidence that neighbourhood geographic scale could be a contributing factor in explaining the inconsistent findings of the association between neighbourhood food environment and individual obesity risk in the literature. However, it is important to note that contradictory findings from past studies could not always be explained away by differences in neighbourhood geographic scale alone. For example, although some past findings were consistent with ours for the level of geographic scale used, such as a ZIP code level analysis for large grocery stores[6] and a Census block group level analysis for fast food restaurants,[18] other findings disagreed with our results for the level of geographic scale used, such as two significant findings for large grocery stores using Census tract.[18 19] It should be noted that neighbourhood geographic scale is only one consideration in understanding the relationship between neighbourhood food environments and individual obesity risk. Past studies varied in several other aspects such as the food outlet definition and classification, inclusion or exclusion of multiple types of food outlets and food outlet data sources. Future research should investigate the role of these other aspects of neighbourhood food environments in the estimated association between neighbourhood food environments and individual obesity risk in order to better understand the complexity of the associations.

The negative association between the number of convenience stores in the neighbourhood and individual obesity risk is contradictory to much of the existing literature, with the exception of Zick et al[17] using Census block group as the geographic scale. Aside from possible regional differences in what kind of food convenience stores might carry, Zick et al[17] argued that neighbourhood convenience stores could increase an individual's energy output if residents walked to these facilities rather than driving to them. In any event, differences in geographic scale could not explain the differences in our finding and previous findings regarding convenience stores as past studies have used variaous geographic scales with mixed results.[11 18 19 40]

Table 4 OR for control variables: tract-level estimates

Effect	Men			Women		
	OR	95% CL	Pr>ltl	OR	95% CL	Pr>ltl
Age	1.027	(1.026, 1.028)	<0.0001	1.035	(1.034, 1.037)	<0.0001
Tract median family income in $1000	0.993	(0.991, 0.996)	<0.0001	0.984	(0.980, 0.987)	<0.0001
Tract (%)African-American	0.961	(0.939, 0.984)	0.0011	0.975	(0.944, 1.007)	0.131
Tract (%) Pacific Islander/Hawaiian	1.026	(1.008, 1.044)	0.0042	1.025	(1.000, 1.051)	0.0481
Tract (%) Hispanic	1.014	(1.010, 1.017)	<0.0001	1.025	(1.020, 1.030)	<0.0001
Tract (%) Asian	0.969	(0.958, 0.981)	<0.0001	0.969	(0.952, 0.985)	0.0002
Tract population density/miles^2 in 1000	0.991	(0.979, 1.003)	0.1297	0.988	(0.972, 1.005)	0.1533
Intersection density 1 km buffer	1.005	(1.003, 1.007)	<0.0001	1.005	(1.003, 1.007)	<0.0001
Tract median housing age in years	0.986	(0.984, 0.988)	<0.0001	0.988	(0.985, 0.990)	<0.0001

Most of our control variables have associations with obesity risk that are consistent with past research.[41] However, the finding that higher street intersection density was linked to higher obesity risk is counter to much of the exisiting literature but consistent with other studies using Salt Lake County data.[17 26 34] It may be that greater street connectivity reflects greater car traffic, which could discourage individuals from walking or biking. However, it could also be realted to Salt Lake County having unusually wide streets and large street blocks that may discourage walking even when street connectivity is high.[17 42] In addition, an increase in the proportion of African-Americans in the neighbourhood was related to lower odds of obesity, which is counter to the picture provided by national statistics.[43] In the 2000 Census for Utah, African-Americans constituted fewer than 1% of the residents and had somewhat higher incomes than their national counterparts.[44] As such, our results may not generalise to neighbourhoods in other parts of the country that have higher concentration of black residents with lower income.

In empirical research addressing the relationship between neighbourhood environments and obesity risk, researchers often have to choose one neighbourhood geographic scale instead of having the luxury of varying geographic scales by food outlet type. Our findings suggest that if the research focus is on large grocery stores, a geographic scale larger than the ZIP code may be needed to allow researchers to detect potential significant relationships. If a buffer method is used, a buffer radius of more than 1 km should be considered. On the other hand, if limited-service restaurants is the focus, then using Census tract (or larger) would be more effective in identifying relationships. If convenience stores and full-service restaurants are the focus, then using Census tract or Census block group would be better. If all four types of outlets are investigated, then our results suggest that Census tract may be better than either Census block group or ZIP code in terms of reaching a compromise among various outlet types. Using Census tract as the neighbourhood geographic scale was also recommended by Krieger et al,[45] who found that tract-level analyses most consistently indicated associations between neighbourhood socioeconomic characteristics and residents' health. The 1 km buffer measure performs as well as the Census tract measure but requires knowledge of individuals' addresses and is much more computationally intensive.

Regardless of the geographic scale, the fairly consistent results for the direction of associations between individual obesity risk and counts of large grocery stores, convenience stores, limited-service restaurants and full-service restaurants across multiple geographic scales lend support to the validity of these relationships, although the particular geographic scale used can lead to different effect sizes and different levels of statistical significance.

We temper our findings with several caveats. First, as with most studies of neighbourhood characteristics, our results are subject to questions about causal inference. If non-random residential selection exists, the magnitude of our food environment estimates could be misstated.[46] Second, self-reported height and weight can systematically underestimate true obesity status. But, as long as such self-reporting bias is not geographically patterned, the multivariate results should not be affected. Third, commercial business directories such as the D&B are subject to significant error rates in listing and in

Table 5 Area of statistical significance (p value <0.05) for different types of food outlets by gender: multivariate results

	Area of significance for men	Area of significance for women
Large grocery stores	Not significant at all four geographic scales	Not significant at all four geographic scales
Convenience stores	Census tract, block group and 1 km buffer	Census tract and 1 km buffer
Limited-service restaurants	Census tract, ZIP code and 1 km buffer	Census tract, ZIP code and 1 km buffer
Full-service restaurants	Census tract, block group and 1 km buffer	Census tract and 1 km buffer

classification.[47–49] Fourth, availability and price of foods can vary dramatically within a store type, and as such, food store typologies can be limited in capturing the impact of food environment on consumer decisions regarding food consumption (Rose D, Bodor N, Swalm C, *et al*. Understanding the economic concepts and characteristics of food access. Washington DC, 2009).[22] Fifth, demographic data at the individual level were limited, leading to the potential for spurious correlations due to insufficient control of individual-level variables. Sixth, we did not include some potentially more precise Geographic Information System (GIS) measures of the neighbourhood food environment such as network distance to the nearest or primary food outlet.[50] While this latter measure is more precise, it requires significant computational time and research costs to construct and any data errors could lead to accuracy problems whereas area measures are more forgiving of such data errors.

Our paper contributes to the literature by providing insights into the role of neighbourhood geographic scale in the relationship between food environments and individual obesity risk. The utilisation of driver license records has the advantage of extensive local coverage with a large sample size, allowing us to test a wide range of definitions of neighbourhood, and to estimate models for men and women separately. Moreover, our findings have policy implications. In recent years, policymakers have begun to argue for novel policies that may alter food environments for residents, such as encouraging the building of new large grocery stores or imposing moratoriums on the building of fast food restaurants in certain neighbourhoods. While such policy efforts may be protective for individual obesity risk, understanding the geographic scale of the neighbourhood is important in identifying the area boundaries and location of these interventions.

Author affiliations
[1]Department of Family and Consumer Studies, University of Utah, Salt Lake City, Utah, USA
[2]Department of Family and Preventive Medicine & Population Science, Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah, USA
[3]Department of Family and Consumer Studies & Population Science, Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah, USA

REFERENCES

1. Flegal KM, Carroll MD, Ogden CL, *et al*. Prevalence and trends in obesity among US adults, 1999–2000. *JAMA* 2002;288:1723–7.
2. Ogden CL, Carroll MD, Kit BK, *et al*. Prevalence of obesity in the United States, 2009–2010. *NCHS Data Brief* 2012;82:1–8.
3. Glanz K, Sallis JF, Saelens BE, *et al*. Healthy nutrition environments: concepts and measures. *Am J Health Promot* 2005;19:330–3.
4. Powell LM, Auld MC, Chaloupka FJ, *et al*. Associations between access to food stores and adolescent body mass index. *Am J Prev Med* 2007;33:S301–7.
5. Morland K, Wing S, Diez Roux A. The contextual effect of the local food environment on residents' diets: the atherosclerosis risk in communities study. *Am J Public Health* 2002;92:1761–7.
6. Lopez RP. Neighborhood risk factors for obesity. *Obesity* 2007;15:2111–19.
7. Fraser LK, Edwards KL, Cade J, *et al*. The geography of fast food outlets: a review. *Int J Environ Res Public Health* 2010;7:2290–308.
8. White House Task Force on Childhood Obesity. *Solving the problem of childhood obesity within a generation—White House Task Force on Childhood Obesity Report to the President*. Washington DC, 2010.
9. Dubowitz T, Ghosh-Dastidar M, Eibner C, *et al*. The women's health initiative: the food environment, neighborhood socioeconomic status, BMI, and blood pressure. *Obesity* 2012;20:862–71.
10. Bodor JN, Rice JC, Farley TA, *et al*. The association between obesity and urban food environments. *J Urban Health* 2010;87:771–81.
11. An R, Sturm R. School and residential neighborhood food environment and diet among California youth. *Am J Prev Med* 2012;42:129–35.
12. Lee H. The role of local food availability in explaining obesity risk among young school-aged children. *Soc Sci Med* 2012;74:1193–203.
13. Boarnet MG, Sarmiento S. Can land-use policy really affect travel behaviour? A study of the link between non-work travel and land-use characteristics. *Urban Stud* 1998;35:1155–69.
14. Openshaw S. A geographical study of scale and aggregation problems in region-building, partitioning and spatial modeling. *Trans Institutes Br Geogr* 1977;NS 2:459–72.
15. Openshaw S. The Modifiable Areal Unit Problem. *(Concepts and Techniques in Modern Geography)*. Norwich: Geo Books, 1983.
16. Sharkey JR, Horel S. Neighborhood socioeconomic deprivation and minority composition are associated with better potential spatial access to the ground-truthed food environment in a large rural area. *J Nutr* 2008;138:620–7.
17. Zick CD, Smith KR, Fan JX, *et al*. Running to the store? The relationship between neighborhood environments and the risk of obesity. *Soc Sci Med* 2009;69:1493–500.
18. Morland K, Evenson K. Obesity prevalence and the local food environment. *Health Place* 2009;15:491–5.
19. Morland K, Roux AVD, Wing S. Supermarkets, other food stores, and obesity—the atherosclerosis risk in communities study. *Am J Prev Med* 2006;30:333–9.
20. Maddock J. The relationship between obesity and the prevalence of fast food restaurants: state-level analysis. *Am J Health Promot* 2004;19:137–43.

21. Bodor JN, Rose D, Farley TA, *et al*. Neighbourhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. *Public Health Nutr* 2008;11:413–20.

22. Rose D, Hutchinson PL, Bodor JN, *et al*. Neighborhood food environments and body mass index: the importance of in-store contents. *Am J Prev Med* 2009;37:214–19.

23. Hutchinson PL, Nicholas Bodor J, Swalm CM, *et al*. Neighbourhood food environments and obesity in southeast Louisiana. *Health Place* 2012;18:854–60.

24. Michimi A, Wimberly MC. Associations of supermarket accessibility with obesity and fruit and vegetable consumption in the conterminous United States. *Int J Health Geogr* 2010;9:49.

25. Powell LM, Han E. The costs of food at home and away from home and consumption patterns among U.S. adolescents. *J Adolesc Health* 2011;48:20–6.

26. Yamada I, Brown BB, Smith KR, *et al*. Mixed land use and obesity: an empirical comparison of alternative land use measures and geographic scales. *Prof Geogr* 2012;64:157–77.

27. The University of Utah Pedigree and Population Resource. Utah Population Database. 2010.

28. Bender R, Jockel KH, Trautner C, *et al*. Effect of age on excess mortality in obesity. *JAMA* 1999;281:1498–504.

29. U.S. Bureau of the Census. Salt Lake County, Utah, Census 2000 Demographic Profile Highlights, Summary File 3. 2000. http://factfinder.census.gov

30. U.S. Bureau of the Census. Census of Population and Housing, 2000: Summary Files 1 and 3. 2000.

31. U.S. Census Bureau. 2006 Second Edition TIGER/Line Files. 2006.

32. Handy S. Urban form and pedestrian choices: a study of four Austin neighborhoods. *Transportation Res Record* 1996;1552:135–44.

33. King WC, Belle SH, Brach JS, *et al*. Objective measures of neighborhood environment and physical activity in older women. *Am J Prev Med* 2005;28:461–9.

34. Smith KR, Brown BB, Yamada I, *et al*. Walkability and body mass index: density, design, and new diversity measures. *Am J Prev Med* 2008;35:237–44.

35. Cervero R, Kockelman K. Travel demand and the 3Ds: density, diversity, and design. *Transport Res Part D Transport Environ* 1997;2:199–219.

36. Dun & Bradstreet. Dun & Bradstreet Database. 2000.

37. Matheson FI, Moineddin R, Glazier RH. The weight of place: a multilevel analysis of gender, neighborhood material deprivation, and body mass index among Canadian adults. *Soc Sci Med* 2008;66:675–90.

38. Ahluwalia IB, Mack KA, Murphy W, *et al*. *State-specific prevalence of selected chronic disease-related characteristics: Behavioral Risk Factor Surveillance System, 2001. MMWR Surveillance Summaries*. Hyattsville, MD: Centers for Disease Control and Prevention, 2003.

39. Lin M, Lucas HC, Shmueli G. Too big to fail: large samples and the p-value problem. *Inf Syst Res* 2013;24:906–17.

40. Gustafson AA, Sharkey J, Samuel-Hodge CD, *et al*. Perceived and objective measures of the food store environment and the association with weight and diet among low-income women in North Carolina. *Public Health Nutr* 2011;14:1032–8.

41. Feng J, Glass TA, Curriero FC, *et al*. The built environment and obesity: a systematic review of the epidemiologic evidence. *Health Place* 2010;16:175–90.

42. Hess PM, Moudon AV, Snyder MC, *et al*. Site design and pedestrian travel. *Transport Res Record* 1999;1674:9–19.

43. Ogden CL, Carroll MD, Curtin LR, *et al*. Prevalence of overweight and obesity in the United States, 1999–2004. *J Am Med Assoc* 2006;295:1549–55.

44. Zick CD, Smith KR. *Utah at the beginning of the new millennium: a demographic perspective*. Salt Lake City, UT: University of Utah Press, 2006.

45. Krieger N, Chen JT, Waterman PD, *et al*. Race/ethnicity, gender, and monitoring socioeconomic gradients in health: a comparison of area-based socioeconomic measures—the public health disparities geocoding project. *Am J Public Health* 2003;93:1655–71.

46. Zick CD, Hansen H, Fan JX, *et al*. Re-visiting the relationship between neighbourhood environment and BMI: an instrumental variables approach to correcting for residential selection bias. *Int J Behavioral Nutr Phys Act* 2013;10:27.

47. Liese AD, Colabianchi N, Lamichhane AP, *et al*. Validation of 3 food outlet databases: completeness and geospatial accuracy in rural and urban food environments. *Am J Epidemiol* 2010;172:1324–33.

48. Powell LM, Han E, Zenk SN, *et al*. Field validation of secondary commercial data sources on the retail food outlet environment in the U.S. *Health Place* 2011;17:1122–31.

49. Gustafson AA, Lewis S, Wilson C, *et al*. Validation of food store environment secondary data source and the role of neighborhood deprivation in Appalachia, Kentucky. *BMC Public Health* 2012;12:688.

50. Drewnowski A, Aggarwal A, Hurvitz PM, *et al*. Obesity and supermarket access: proximity or price? *Am J Public Health* 2012;102:e74–80.