A systematic review employing the GeoFERN framework to examine methods, reporting quality and associations between the Retail Food Environment and obesity

Abstract

This systematic review quantifies methods used to measure the 'retail food environment' (RFE), appraises the quality of methodological reporting, and examines associations with obesity, accounting for differences in methods. Only spatial measures of the RFE, such as food outlet proximity were included. Across the 113 included studies, methods for measuring the RFE were extremely diverse, yet reporting of methods was poor (average reporting quality score: 58.6%). Null associations dominated across all measurement methods, comprising 76.0% of 1937 associations in total. Outcomes varied across measurement methods (e.g. narrow definitions of 'supermarket': 20.7% negative associations vs 1.7% positive; broad definitions of 'supermarket': 9.0% negative associations vs 10.4% positive). Researchers should report methods more clearly, and should articulate findings in the context of the measurement methods employed.

Key Words: Community Nutrition Environment, Foodscape, GIS, Food Access, Obesogenic Environment, Methodology

Introduction

The idea that the retail food environment (RFE) is a cause of obesity is intuitively appealing. The RFE comprises the spatial availability, accessibility and composition of food outlets within local environments. Over the past half-century, the RFE has changed drastically. Since the 1970s many western countries have seen a shift in grocery retailing with large superstores establishing in suburban and out of town regions, leading smaller, local high-street grocers to close (1, 2, 3, 4). This has purportedly led to the development of so-called 'food deserts', where residents lack access to healthy food. Evidence supporting existence of food deserts is particularly strong within the US (5, 6, 7). The US and UK have also seen a proliferation of restaurants and fast food outlets, providing low-cost, energy-dense foods (8, 9, 10). These changes have coincided with increases in obesity rates, which have been rising globally since the 1970-80s (11). In the UK (12, 13, 14), and internationally (15, 16) policymakers have been investigating ways to intervene to create healthier RFEs, for example through banning the opening of fast food outlets around schools (17).

Despite considerable research activity, evidence supporting a link between the RFE and obesity is conflicting at best. The largest systematic review to date on the association between the RFE and weight status (18) in the US and Canada found for example that while there were 31 statistically significant positive associations between fast food outlets and increased weight status, 99 associations were null, and 4 were negative (showing increased number or proximity of fast food outlets was associated with decreased weight status). Similarly, supermarkets (often considered as a proxy for healthy food availability) were found to be statistically significantly positively associated with lower weight status in only 24 of 143 associations, with 7 associations in the unexpected direction (showing increased number or proximity of supermarket access was associated with higher weight status). Other systematic reviews have reported predominantly null findings in relation to RFE-obesity associations (19, 20, 21, 22). Although these reviews focus only on p-values, and do not account for the magnitude of associations, they do highlight numerous conflicting results and tend to suggest little or no consistent link between the RFE and obesity.

A common challenge in understanding RFE-obesity associations - repeatedly noted by authors – is the diversity of methods used to measure the RFE (18, 19, 20, 21, 22). The majority of literature uses spatial measures, such as the density or proximity of outlets, to characterise the RFE (23, 24). A recent review identified five dimensions of methodological diversity with regard to spatial RFE measures: (i) source of food environment data, (ii) methods used to extract food outlets from a wider dataset, (iii) methods for classifying outlets, (iv) geocoding methods, and (v) choice of RFE metric or measure (25). These are summarised in the GeoFERN reporting framework: a reporting checklist developed specifically for RFE research covering the five dimensions. A number of reviews have quantified methods used in RFE literature across aspects of these domains (18, 20, 21, 26). However, no study has systematically and comprehensively quantified the degree of methodological diversity across all five domains. Quantification of the methods used is important to (i) identify priority areas for future research into the impact of methods, and (ii) highlight the scale of the problem in order to encourage researchers to move towards more standardised or evidence-based methods where possible. One aim of this review was therefore to

(1) **conduct a systematic review to comprehensively identify the spatial methods used to measure the RFE across the five GeoFERN dimensions and quantify their frequency of use.**

In this context, we were interested in spatial measures that consider the locations of all food outlets (of a given type) within a neighbourhood. Thus studies that measure specific outlets in isolation, such as the presence of a newly opened supermarket, are not considered in this review. This is because these measures do not consider the wider food environment (e.g. the availability of pre-existing supermarkets within the neighbourhood), and thus do not provide standardised measures that would be translatable across environments.

Previous work (25) has also highlighted that methodological information is often not reported in papers. However, no study has ever quantified the quality of methodological reporting, and thus little is known about the scale of the problem of incomplete reporting. A second aim was therefore to:

(2) Quantify the quality of methodological reporting within the spatial RFE literature.

Given the varied approaches to measuring the RFE, it is perhaps not surprising that evidence is conflicting. Recent research suggests that methods used to measure the RFE may impact RFE-obesity relationships. For example, while many studies operationalise the RFE in terms of singular food outlet types, such as ‘fast food outlets’ (18, 20, 21), research suggests that relative measures of food outlet mix (e.g. the ratio of fast food outlets to supermarkets) may be more strongly and consistently associated with obesity-related outcomes (27, 28, 29, 30). Other methodological factors such as choice of food environment data (31, 32) and geocoding methods (33) are also beginning to be investigated.

Associations between the RFE and obesity may additionally vary across population groups. Stronger associations between the RFE and obesity-related outcomes have been found within more deprived neighbourhoods (34, 35, 36). Differential associations have also been observed for people of differing income and education (37, 38), ethnicity (39), age (40) and across urban/rural residences (34).

Existing systematic reviews either do not account for potential divergent effects across measurement methods or population groups (20, 22), or account only for a limited range of factors using simplistic groupings of studies; for example grouping diverse methods together (18, 19, 21, 23). These reviews may therefore miss important associations at the level of population groups or measurement methods. A final aim of this systematic review was therefore to:

(3) Examine the evidence for associations between the spatially measured RFE and obesity, accounting for possible divergent associations across measurement methodologies.

Methodology

Data Sources and Search Strategy

This review capitalises on work carried out by Cobb *et al.* (18) by updating and expanding upon their existing systematic review into associations between the RFE and obesity. Papers identified by Cobb *et al.* (18) ('the Cobb review') are included in the present review (subject to exclusion criteria), and were re-reviewed to extract new information as outlined below. Additionally, the Cobb search was re-run to identify latterly published literature, including other western countries (signatories to the Organisation for Economic Co-operation and Development convention) in addition to the US and Canada. Non-western countries were excluded from the present review due to differences in food environment and obesity dynamics (41).

Mirroring the Cobb review, searches were performed using Scopus and PubMed for English-language literature published online or in print relating to the association between the RFE and obesity, using search terms alluding to weight status, such as ‘overweight’, ‘obese’ and ‘body mass index’ and to the RFE, such as ‘food environment’, ‘food access’ and ‘fast food’ (Supplement 1). We sought to identify literature published from the 1st January 2014 (to align with the end of the Cobb review) up to the 8th June 2017. To capture literature that was published but not indexed before 1st January 2014, we searched by ‘date created’ on PubMed, and allowed a 1-year time-lag for Scopus; thus including in our Scopus search all literature published since 1st January 2013.

Exclusion Criteria

All exclusion criteria replicated those of the Cobb review unless otherwise indicated. More particularly, in accordance with the Cobb review, studies in our review were required to examine associations between objective spatial measures of the RFE around the home and individual-level outcomes of either BMI, weight classification (e.g. ‘obesity’), BMI change or weight change (referred to collectively as measures of ‘obesity’). Replicating Cobb, our review focussed on home environments, which are the most commonly investigated environments (21). Further following Cobb, studies in the present review were excluded if they (i) examined associations with area-level outcomes only (e.g. obesity prevalence), (ii) treated the RFE as a moderator, mediator or covariate only, (iii) used simulated data, (iv) combined RFE measures with other environmental measures (e.g. access to physical activity facilities), such that the effects of the RFE could not be isolated or (v) if they were case studies investigating the influence of one or more specific outlets, such as a newly opened supermarket or a store that a participant has visited, without measuring the wider RFE. In line with the

Cobb review, studies in the present review were additionally required to (i) include at least 200 people, (ii) operationalise the RFE using areal units smaller than or equal to US zip code zones, and (iii) examine spatial measures of at least one of: supermarkets, grocery stores, convenience stores, fast food restaurants, full-service restaurants, composite measures including at least one of these outlet types, or food availability within at least one of these outlet types. One objective of the present review was to evaluate the reporting of methods, and thus, as a departure from the Cobb strategy, papers were required to report primary research within peer-reviewed journals.

Screening

The top-up search returned 4,801 results, which were exported to Endnote for deduplication. Of the remaining results (n = 3,984), 1,844 articles were excluded after title screening; 1,776 after abstract screening and 317 after full-text screening. Five studies from the Cobb review were excluded in line with our exclusion criteria (Figure 1). Overall 113 papers were included in the review.

All articles were screened by the primary reviewer (EW). Double-screening was undertaken for a sample of articles (2015 titles; 1276 abstracts; 70 full-texts) by one of six reviewers (CG, DR, MM, MH, WM or AM). Conservatively, reviewers excluded articles at the title stage only if they very clearly were off-topic or met an exclusion criterion and titles were retained if at least one reviewer determined not to exclude them. Disagreements at the abstract and full-text stage were resolved by a third independent reviewer. After full-text screening, agreement with the primary reviewer's decision was 98.6%, with one paper excluded by the primary reviewer being retained.

Data extraction

Data were extracted from both the newly identified studies and the original Cobb studies on study design, RFE measurement methods, study quality, and numbers of null and statistically significant associations, and the directions of statistically significant associations. Data extraction was considerably more extensive than in the Cobb review, totalling 99 data fields (Supplement 2). Methodological information was extracted for each of the reporting items deemed 'essential' in the 'GeoFERN' checklist (25). Effect sizes were not extracted due to the varied methods and measures used, making collation of these data at the scale of the present review impossible. This approach of counting null and significant associations has been employed by other systematic reviews when faced with similarly methodologically diverse data (18, 20, 42). All data were extracted into Microsoft Excel.

Often papers report associations for multiple statistical models, or repeat analyses for different population groups or exposure measures. Outcomes data (numbers of statistically significant/null results) were extracted for each distinct exposure measure, outcome (e.g. BMI or 'obesity') and population subgroup, including results for the full sample, if reported. This is because these different models represent different research questions. Where multiple models were run for the same exposure-outcome-population grouping (e.g. using different covariates), results were only extracted for the 'main model' (Supplement 3). For most studies, the main model was taken to be the most fully-adjusted model presented in a results table.

Due to the scale of the review, only the aims and objectives, methods and results sections of papers were reviewed, except where explicit reference was made to methodological details provided in supplementary materials or other published papers.

Authors were not contacted to obtain missing information due to the high prevalence of missing information.

Data extraction was performed by the primary reviewer. Two second reviewers (MH, AC) independently extracted data from a random 20% sample (n = 23), with disagreements being resolved through discussion. Overall, 96.5% of data fields (n = 2,427) were in agreement with the first reviewer's initial decision (Supplement 4).

Quality screening

Studies were appraised for risk of bias using an expanded version of the Cobb review quality checklist, adapted from the Newcastle Ottawa Scale (43). A total of 10 marks were available for features such as validation of food environment data, use of a causal framework, use of multi-level modelling or equivalent methods accounting for clustering within neighbourhoods (where relevant) and controlling for key covariates (age, race, gender and neighbourhood socioeconomic status/racial composition) (Supplement 5). Quality scores were expressed as a percentage of eligible marks, with higher scores indicating lower risk of bias. Additionally, the completeness of methodological reporting was appraised using the GeoFERN reporting checklist (25). For each paper, one mark was awarded for each 'essential' reporting item on the GeoFERN checklist, with half-marks being awarded when reporting criteria were partially met. An overall GeoFERN reporting score for each paper was calculated as the percentage of eligible marks obtained. After double-screening a 20% sample of papers (n = 23), agreement between the final decision and the first reviewer's initial decision was 96.3% for the GeoFERN scores and 97.0% for the study quality tool.

Data synthesis

We reported the frequency of use of different RFE measurement methods across the five GeoFERN domains (data source, data extraction, food outlet classifications, geocoding methods and RFE metrics) and the prevalence of missing methodological information within each domain. The numbers of null and statistically significant positive/negative associations were reported for 112 studies; one study (44) was excluded from this aspect of the analyses because it did not report the main effects of the RFE. For the four main exposures of 'fast food outlets', 'supermarkets/grocery stores', 'convenience stores' and 'restaurants', results were stratified by population groups, and for the two most common exposures ('fast food outlets' and 'supermarkets/grocery stores'), results were stratified by measurement method. We additionally evaluated the numbers of null and statistically significant positive/negative associations for studies within the top decile of quality score only (≥66.7%), to determine whether our findings were sensitive to study quality. Data were presented for populations and methods used in 5 or more studies to enable generalised comparisons between methods. Further information on the coding of data is available in Supplement 6.

Results

Study Characteristics

There were 113 papers included in this review, published between 2004-2017 (Supplement 7), comprising 107 unique datasets. Sixty-six were identified from the original Cobb review, with the remaining 47 newly identified. The median participant sample size was 3,786 (range: 219 to 453,927). Twelve studies derived outcome data from a dataset that was also used in another study (6 unique datasets). Due to the large number of studies included in this review, only summary data are provided in the main text. However, Supplements 8 and 9 respectively provide detailed information on study characteristics and findings at the level of individual papers.

Descriptive statistics of the studies are presented in Table 1. Overall, studies predominantly related to the RFE in the US (82.3%), examined populations of mixed gender (88.5%), who were adults (66.4%), of mixed races (64.5%) and mixed socioeconomic status (SES) (62.8%). Of those studies reporting environmental context, the vast majority were either mixed urbanity or entirely urban (95.5%). Nearly three quarters of the studies were cross-sectional. Contrasting the newly-identified papers to the older papers from the Cobb review, there was a higher proportion of longitudinal studies (34.0% vs 15.2%), studies relating to ethnic minority populations (12.8% vs 6.1%) and studies in predominantly urban areas (25.5% vs 1.5%). Despite the wider geographic scope of the updated review, a high proportion of studies originated from the US (72.3%).

Study Quality

Study quality scores ranged from 22.2% (indicating a high risk of bias) to 88.9% (indicating a lower risk of bias), with a mean of 49.9%. There were 18 studies (15.9%) with a score ≥ 66.7%, corresponding to the top decile of quality scores. The most common risks of bias were failure to use a causal framework to guide model development (97.3% of studies), failure to control for neighbourhood self-selection (85.8% of studies), and use of secondary food environment data without validation of the data (78.8% of studies). Further data on study quality is presented at Supplement 5, Table S5.

Methods used in the research

Figure 2 displays the frequency of use of different methodologies across studies. Restricting analyses to only those identified in the top-up search did not lead to substantively different findings.

Data Source

The vast majority of studies (83.2%) obtained RFE data from a single source, with the remainder combining multiple sources. Commercial data (for business marketing or market research purposes) and government data were the most common data sources (Figure 2). Commercial data were typically from InfoUSA (InfoGroup, Inc.) (36.2%) or Dun & Bradstreet, Inc. (34.1%) (including the National Establishment Time Series dataset, which is derived from Dun & Bradstreet). Government data were typically from local health, hygiene or licensing departments (71.4%).

Data Extraction

Once RFE data have been obtained, it is often necessary to extract specific food outlets of interest from a wider dataset. Data were predominantly extracted using information within the RFE data, which included proprietary classifications, store names, or other attributes, such as store size or revenue. Some studies used secondary data sources such as business directories and websites (Figure 2). The majority of studies (73.8%) used a single method (e.g. proprietary classifications only), with the remainder using a combination of methods (e.g. proprietary classifications and store names).

Food Outlet Constructs

Studies typically employed ‘fast food outlets’ constructs (sometimes referred to as ‘takeaways’ or ‘limited service restaurants’), ‘supermarkets’ and/or ‘grocery stores’ (hereinafter ‘supermarkets/grocery stores’), ‘convenience stores’ (including ‘bodegas’), and ‘full-service’ or ‘sit-down’ restaurants’ (hereinafter ‘restaurants’) (Figure 2). These outlet constructs were either measured in isolation (for example as the density or proximity of ‘fast food outlets’) or as part of a composite variable, such as the ratio of ‘fast food outlets’ to ‘restaurants’. Forty studies (35.4%) used other food outlet constructs, such as ‘food stores’ or ‘total restaurants’, which encompassed, but did not directly define the four main outlet types. Supermarkets and grocery stores were grouped under one category because studies defined these constructs inconsistently. For example, some studies would use the term ‘grocery store’ to refer to both large chain supermarkets as well as smaller local grocery stores, whereas other studies would treat large chain ‘supermarkets’ and smaller ‘grocery stores’ as distinct constructs.

Constructs were defined using four main methods: (i) use of proprietary classifications within the RFE data, (ii) use of other attributes within the RFE data, such as store name or size, (iii) a combination of proprietary classifications and other attributes within the RFE data, and (iii) telephone or in-person audits (Figure 2). Other methods included use of supplementary information, such as internet searching.

The scope of commonly employed food outlet constructs also varied (Figure 2). For example, 35.2% of studies defined ‘supermarkets’ narrowly to include only large chain supermarkets, 40.7% employed a moderate scope including large/mid-sized grocery stores, and 24.1% included small grocery stores (see Supplement 6 for further details). While several studies cited use of standardised classification schemes (NAICS classification scheme: 23.9% of studies; Standard Industry Classification (SIC) scheme: 13.3% of studies), these were not necessarily employed in the same way. For example, while some studies used the NAICS code 722513 for ‘limited service restaurants' to define ‘fast food outlets’ (45, 46, 47, 48); others additionally included cafeterias (NAICS code 722212) and mobile food services (NAICS code 722330) (49) or pizza restaurants (NAICS code 722211) (50).

Geocoding

Geocoding is the process of converting address information into coordinates or other geographic identifiers through matching of address information to spatially coded reference data. Home addresses were most commonly geocoded to geographic identifiers at the level of census tracts, postcode zones or street segments (Figure 2), with the latter method typically using building numbers to estimate how far along the street an address is located. Less commonly, addresses were geocoded to the building level, zip code level, census blocks or land/tax parcels. Similar methods were used for food outlets (data not presented due to small number of studies (n = 15) reporting this information).

RFE Metrics

The metrics used to measure the RFE predominantly included: (i) buffer metrics assessing the RFE within a given distance of the home, (ii) areal metrics assessing the RFE within a predefined areal unit such as a census tract or zip code zone, (iii) proximity metrics, which measure the distance between the home and one or more food outlets, and (iv) gravity metrics, which effectively combine proximity and buffer metrics by measuring the count or density of food outlets within a buffer, with outlets that are more proximal to the home being weighted higher (Figure 2).

Within these broad types of metric, specific measures were highly diverse. This was particularly true for areal and buffer metrics, as enumerated in Table 2. Areal and buffer measures were used 242 times across the 113 papers. Of the metrics that had a clearly defined unit of measurement, the most common was the count of outlets per unit area, which included counts of outlets within Euclidian (straight-line) buffers (31.2% of measures), followed by raw, non-standardised counts of outlets (19.4% of measures) and presence/absence of an outlet type (14.8% of measures).

The geographic scope of area-based measures also varied. For areal metrics (58 measures), studies most commonly used census tracts to define the scope of the RFE (53.6% of areal measures). Buffer metrics were typically delineated in terms of Euclidian distances (83 measures, 52.2% of buffer metrics) or network distances (50 measures, 31.4% of buffer measures), with 27 measures (17.0%) of undefined delineation. The scope of buffers varied, but were generally between 400-1600m for both network and Euclidian buffers (59.7% of buffer measures). Nearly half of all studies that employed buffer metrics (46.8% of 62 studies) measured the same RFE metric using more than one buffer size. Seven studies included 2 buffer sizes, 10 investigated 3 buffer sizes, and 12 investigated 4 or more buffer sizes.

Proximity measures were used 36 times across the 113 studies. These metrics were also variable, but to a lesser extent than for area-based metrics, with the vast majority (88.9%) measuring the distance to the nearest outlet of a given type e.g. ‘supermarket’. Alternative proximity measures included the average distance to the nearest ‘N’ outlets of a given type (5.6% of measures), and the relative proximity of two or more outlet types, such as the distance to the nearest healthy outlet minus the distance to the nearest unhealthy outlet (5.6% of measures). Proximity was most commonly measured as the network distance (93.5% of measures), with Euclidian distance and travel time being used an equal number of times (16.1% of measures respectively).

Gravity metrics were also varied. Of the five studies that used gravity measures (4.4%), four of these used a fixed circular bandwidth, which ranged from 1 km to 6 miles and one used an adaptive bandwidth, but it was unclear what the adaptive radius was based on. Two studies used a quartic decay function (defining how quickly the weighting of food outlets falls off with increasing distance), one used a quadratic, one used a Gaussian decay function and one did not report the decay function used.

Quality of Methodological Reporting

Overall, the mean GeoFERN reporting quality score was 58.6% (range: 25.0%-97.2%). Table 3 shows the completeness of methodological reporting across the five GeoFERN domains. Methodological reporting was worst for the geocoding domain, with only 3 papers (2.7%) providing full information on the geocoding methods used and an average score of 41.2%. It was commonly unclear whether geocoding was used and/or how this was performed, with this information being omitted in relation to the geocoding of food outlets and homes in 76.1% and 58.4% of studies respectively.

RFE-Obesity Associations

Overall, there were 1,937 reported associations between the RFE and obesity. Null associations predominated, making up 76.0% of all associations. Table 4 enumerates the associations between the most common measures of the RFE (fast food outlets, convenience stores, supermarkets/grocery stores, and restaurants) and obesity, including sub-groups of age, gender, ethnicity, and urban/rural status. Throughout, ‘positive associations’ refer to statistically significant associations indicating increased access/exposure to food outlets is associated with *increased* obesity, and ‘negative associations’ refer to statistically significant associations indicating increased access/exposure to food outlets is associated with *decreased* obesity.

The distribution of associations varied across population groups. For example, there was a stronger tendency toward more positive associations than negative associations for fast food outlets among low-SES children (39.3% positive, 3.6% negative, 57.1% null) than for the general population (20.8% positive, 4.2% negative, 75.0% null). Additionally, there was no trend towards positive/negative associations for convenience stores among the general population (10.9% positive, 9.5% negative, 79.6% null). However, after stratifying by age, convenience stores tended to be more consistently associated with higher rather than lower obesity among children (16.5% positive, 2.6% negative, 80.9% null); particularly those of low-SES and non-white ethnicity (e.g. 39.3% positive, 3.6% negative, 57.1% null for low-SES). Restricting to high-quality studies did not substantively change findings.

Table 5 shows the distribution of positive, negative and null associations after stratification by RFE measurement method. Results by geocoding method are not presented, because geocoding methods were rarely reported. The distribution of positive and negative associations differed across RFE measurement methods. In particular, there was very little evidence supporting an association between supermarkets and obesity when considering all definitions of ‘supermarket’ collectively (6.6% positive, 12.9% negative, 80.5% null). However, when considering only narrow definitions of ‘supermarket’ (i.e. only large chain outlets), there was a tendency for more negative than positive associations (1.7% positive, 20.7% negative, 77.6% null). Additionally, there was a tendency for more positive than negative associations for narrowly defined measures of fast food outlets (major chain outlets only) compared to broader definitions (e.g. 26.1% of associations vs 19.4%).

The distribution of associations additionally varied across RFE metrics. For example, there was a tendency for more positive associations for measures of count/area, count/population and proximity of fast food outlets than for measures of presence/absence (e.g. proximity of fast food outlets: 28.6% positive, 2.6% negative, 68.8% null; presence/absence of fast food outlets: 0% positive, 3.8% negative, 96.2% null) and raw, non-standardised count (18.9% positive, 13.5% negative, 67.6% null). Measures of relative unhealthiness (such as the ratio of fast food outlets to total outlets) also tended notably towards more positive than negative associations (21.3% positive, 0% negative, 78.7% null). Additionally, there was a stronger tendency towards positive associations for fast food outlets among children for buffers ≤400m (25% positive, 0% negative, 75% null) than for larger buffers (2.0% positive, 6.1% negative, 91.8% null). Finally, use of commercial data tended to be associated with a stronger patterning of associations in the expected directions for both fast food outlets and supermarkets.

Discussion

Methodological Diversity and Reporting

Existing systematic reviews into the RFE and obesity have repeatedly noted the diversity of methods used to measure the RFE (18, 20), often pointing to this diversity as limiting or even precluding conclusions that can be drawn from the evidence base (19, 21, 22). However, no review has ever comprehensively quantified the diversity of methods across all aspects of methodological diversity, and thus the scale of this problem is unknown. This study extends the evidence base by systematically quantifying methods used across the five dimensions of methodological diversity: (i) the source of food environment data, (i) the methods used to extract food outlets from a wider dataset, (iii) the methods and definitions used to classify outlets, (iv) geocoding methods and (v) RFE metrics, including all important methodological details rated as 'essential' in the GeoFERN framework (25). Understanding the methods used in the RFE literature will support emerging research into the comparability of different methods, by highlighting priority areas for further research. This review also quantifies for the first time the prevalence of missing methodological information relating to measurement of the RFE. Methodological information is critical to the interpretation of RFE-obesity studies, particularly given the mixed methods employed, and thus awareness of the extent of the issue will help motivate improved reporting moving forward.

A key finding of this review was that the degree of methodological diversity was extremely high. This finding is in agreement with the earlier Cobb review, who also found considerable methodological diversity in the literature. However, our review provides further information on the methods used across all dimensions of methodological diversity and across a wider selection of countries. In particular, our review provides new information on the methods used to extract food outlets from secondary datasets, apply food outlet classifications, and geocode food outlet and participant addresses. It also quantifies for the first time the variability in food outlet classification scopes and elucidates the true scale of diversity of areal measures of the RFE, which differ both in relation to their scope and unit of measurement. This diversity makes the collation and interpretation of research very challenging, as little is known about the comparability of different methods.

A second key finding was that RFE measurement methods are not well reported in the literature. Indeed, we found that not one single study provided all details rated as ‘essential’ within the GeoFERN reporting framework, and 33 studies (29.2%) provided less than half of these details. Overall, the high prevalence of missing methodological information, combined with the diversity of methods, severely limits the inferences that can currently be drawn from the evidence base. While policymakers should be praised for taking action against potentially obesogenic RFEs, inadequate methodological reporting undermines these efforts. We suggest that authors and journal editors take greater responsibility for ensuring the complete reporting of RFE measurement methods, for example through use of the GeoFERN framework (25). A reduction in the diversity of measures used would also be of benefit moving forward. Researchers should give closer scrutiny to the methods used to ensure, where possible, that the best or most accurate methods are being used, such as use of validated secondary RFE data and accurate geocoding methods.

It is hoped that the findings from this review motivate further research into the comparability of methods within each of the five dimensions of diversity. Some research has been done relating to the choice of food outlet data (31, 32, 51, 52, 53, 54) and RFE measures (27, 28, 29, 30, 50, 55). However, the other dimensions remain largely unexplored. Understanding the impacts of different methodological approaches will not only aid collation and interpretation of existing research, but may highlight best practice methods and help standardise measures used in future research.

One particular priority area for future research is in relation to the definition of food outlet constructs. Considerable diversity was observed across food outlet definitions. For example, fast food outlets were often defined narrowly as comprising only chain fast food outlets, and in other cases were defined broadly to include not only traditional non-chain fast food outlets, but also outlets such as coffee and sandwich shops, and desert shops. This diversity exists in spite of the existence and frequent citation of several industry-standard classification schemes (NAICS and SIC). Indeed, even when standardised classification schemes were cited, they were inconsistently applied. To our knowledge, no study has ever explored the impact of using different definitions for a given outlet construct, so it is unclear whether distinctions between different definitions of outlet constructs are important.

One dimension with particularly high diversity was the choice of RFE measure. For example, while areal and buffer metrics were used 242 times across the 113 studies, specific measures were used, at most, 15 times (count per area within 800m - 1,600m

Euclidian buffers), and commonly no more than once. As mentioned, some research is beginning to investigate the impacts of using different measures – often focussing on the difference between ‘relative’ (e.g. ratio of healthy to unhealthy outlets) and ‘absolute’ (e.g. outlet count) measures (27, 29, 56) or buffer sizes (33, 57, 58). However, given the high degree of diversity among RFE metrics, this remains another key area for further research.

RFE-obesity associations

Previous reviews of RFE-obesity associations are limited in that they do not account for differences in measurement methods when collating the evidence (20, 22), or only account for methods in relatively simplistic ways (18, 21). However, collation of evidence from disparate methods may be misleading and could hide important associations. This review is the first to systematically stratify study findings by detailed methodological characteristics in order to examine how these factors may interact with outcomes. While reporting of methods was generally poor, there was a sufficient number of studies reporting methodological information with each domain to enable comparisons across methods; with the exception of the geocoding domain.

In agreement with existing reviews (19, 20, 21, 22), we found that overall, null associations considerably outnumbered statistically significant associations. This review is the first to demonstrate, however, that null associations remain the dominant outcome across all RFE measurement methods. The impact on effect sizes was not considered due to the diverse methods, which made collation of effect sizes impossible at the scale of this review. However, the high prevalence of null results does suggest any associations are likely to be small, irrespective of the methods used, given the large sample sizes used in most studies. That said, there was a tendency toward more positive than negative associations for fast food outlets, which persisted across most methods (for 18/22 investigated methodological groupings, positive associations were between 16-36% of all associations while negative associations were <5%). As p-values are a function of sample size, these findings do not imply meaningfulness of an association. Nevertheless, a consistent trend towards more associations in one direction versus another may be suggestive of a 'true' association; albeit of unknown magnitude and recognising that these trends might be an artefact of publication bias, or diversity across methods and populations. Additionally, the influence of methods and population characteristics on the distribution of statistically significant associations is of interest in itself, given that p values seem to be the key outcome many authors and policymakers focus on (59).

A further key finding was that the distribution of null and statistically significant associations varied across measurement methods. While it is not possible to attribute this variation entirely to methodological factors (due to differences across studies in sample size, context and other methodological factors not accounted for within methodological groupings, or simply by chance), there were some notable differences that warrant further investigation. Researchers should also ensure that findings are interpreted in view of the methods employed; particularly when collating evidence and translating research into policy.

Of particular note, the distributions of positive, negative and null associations were more supportive of RFE-obesity associations for narrower definitions of 'supermarkets' compared to broad definitions. This is a novel finding; as mentioned above, no study has investigated the impact of construct definitions on associations with obesity. Theoretically, narrow construct definitions may provide better measures of the RFE as they may capture food outlets with a more consistent type of food provision. These findings reinforce the above-mentioned need for research into the comparability of different construct definitions and for researchers to clearly define food outlet constructs.

We additionally found the distribution of associations varied across different RFE metrics. For example, the tendency towards positive rather than negative associations between fast food outlets and obesity was considerably stronger for proximity measures than for measures of presence/absence (e.g. 28.6% positive, 2.5% negative vs 0.0% positive, 3.8% negative). Of relevance to RFE policy, which often restricts development of new fast food outlets within 400m of schools (12, 13), the distribution of associations was more strongly supportive of a link between fast food outlets and obesity among children for buffers ≤400m than for larger buffer sizes. These findings are in broad support of other newly emerging research, which has shown different RFE metrics may critically impact the strength and direction of associations observed between the RFE and obesity-related outcomes, both in terms of the type/unit of measurement (27, 28, 30, 55, 60, 61), and the geographic scope (33, 57, 58).

It is also worth noting that the distribution of associations varied across population groups. ‘Deprivation amplification’ – whereby people of lower-SES are more strongly influenced by their immediate RFE – has been observed in several studies (37, 62), and we found a stronger tendency toward more positive than negative associations for convenience stores and fast food outlets among low-SES groups. In spite of this, many studies do not allow for potential divergent effects across population groups (possibly due to insufficient sample sizes), potentially hiding important associations and explaining the high prevalence of null results. If policymakers are to intervene in relation to the RFE, it is imperative that we understand interactions between the RFE and population characteristics to ensure that interventions do not lead to widening health inequalities.

Limitations of existing research

The studies included within this review had several limitations in addition to those noted above. Overall, study quality was relatively poor, suggesting many studies are at risk of bias. Of most concern, given that this evidence is often used to inform RFE interventions, was the absence of causal frameworks from all but three studies. Causal frameworks inform covariate selection to allow more robust causal inference in observational research (63). In relation to this, many studies did not account for competing aspects of the built environment which may be correlated with RFE measures. For example, places that have a high availability of unhealthy food retailing may also have a high availability of healthy food retailing, and may be more conducive of walking, due to a higher accessibility of general facilities/destinations (28, 64). Without accounting for such competing exposures, associations between specific RFE measures and obesity will be biased. Use of a causal framework would identify these necessary covariates. Recent evidence also suggests that mutual adjustment for competing food outlet types (e.g. both 'healthy' and 'unhealthy' outlets) may be critical in detecting statistically significant associations (27, 35, 65, 66), although many of these studies often found no appreciable impact on effect sizes. The above notwithstanding, we found no substantive differences in our findings after restricting to papers within the top decile of quality score.

This review also highlights the vast number of studies that have examined the RFE around the home. However, GPS and travel diary studies show that home-centric neighbourhoods do not correspond well with people's actual activity spaces (67, 68), raising questions around the appropriateness of home-centric measures. It is also notable that the majority of research – including numerous longitudinal studies – measure the RFE at only a single time-point, limiting causal inferences. Studies investigating changes in the RFE through the 1970s - 1990s, when the RFE saw the greatest shifts in food retailing (1, 2, 3, 4) may provide the greatest opportunities for understanding the impact of the RFE on weight status. Further limitations of the RFE-obesity literature include lack of data on food outlet utilisation and the within-store environment (e.g. pricing, food quality) and failure to account for alternative purchasing opportunities, such as online supermarkets, delivery services, and non-traditional food stores, such as clothes shops and pharmacies (69). Many of these limitations appear to be driven by the availability of secondary data (or lack thereof). Nevertheless, use of spatial methods to operationalise the RFE can also be celebrated in that it has enabled investigation of the RFE at a population level; which is important for national and regional-level policymaking.

Strengths and Limitations of Review

This review has several strengths, most notably our systematic search strategy, the very large number of studies included in the review, and the breadth and detail of the data extraction, which provide rich information on the methods used and allow detailed analysis of the distribution of RFE-obesity associations, accounting for measurement methods and population groups.

It is worth reiterating that we decided *a-priori* not to extract effect sizes, because the heterogeneity of RFE measures would preclude collation of these data. Following similar approaches to other reviews in this area (18, 20, 42), we instead counted the distribution of null and statistically significant associations, together with their associated directions. Our findings do not provide any information regarding the strength of associations. Indeed, the p-value is a function of sample size, and the presence/absence of a significant p-value thus does not imply meaningfulness of an association. Nevertheless, by collating the numbers of statistically significant associations across multiple studies, we were able to infer the possible presence of ‘true’ associations (of unknown size) from the distribution of associations. In the absence of any ‘true’ association, the numbers of spurious statistically significant positive and negative associations should be approximately equal. The greater the tendency for more statistically significant associations in one direction than the other, the more suggestive the data of a ‘true’ association. A limitation of this approach is that publication bias may tend to inflate the numbers of associations in the expected direction. That said, positive and negative associations were balanced for supermarkets/grocery stores, convenience stores and restaurants across the general population, suggesting our results may not be substantively impacted by publication bias. Nevertheless, our findings need to be interpreted with caution in this regard.

This review is the first to consider in detail the methods used to measure the RFE when collating the evidence base. However, within methodological groupings, there was still heterogeneity, which may have confounded our results, and it is not possible to attribute variation in the distribution of associations to methodological factors alone. Due to the high prevalence of missing methodological information, we did not contact authors to obtain missing data, and our results are therefore only representative of those studies that reported complete information for a given methodological aspect. We reviewed the aims and objectives, methods and results sections of papers in detail, so may have occasionally missed methodological information that was reported elsewhere. The Cobb review was limited to studies conducted in the US and Canada only. While we expanded the top-up search to cover other western countries, reliance on the Cobb review to identify earlier studies means US and Canadian studies are over-represented, reducing the generalisability of our findings across western countries. Nevertheless, in sensitivity analyses we restricted our analyses to only those studies identified in the top-up search, and found no notable differences as compared to the full sample of studies. This is unsurprising, given the dominance of US studies both in the Cobb review and the top-up search, suggesting our findings are still of reasonable generalisability across western countries. Lastly, we did not consider other differences across studies such as analysis methods or outcome measures, which may have also impacted study findings.

Conclusion

Associations between the RFE and obesity are nuanced, and depend upon the methods used to measure the RFE. However, null associations appear to be the predominant outcome across all measurement methods. At present, the reporting of methods is poor, and severely limits inferences that can be drawn from the evidence base, and the translation of evidence into policy. Authors and journal editors should ensure more robust reporting of RFE measurement methods, for example through use of specially developed reporting frameworks. Authors are also responsible for articulating study findings in the context of the methods employed, so that policymakers can correctly interpret RFE-obesity associations. Direct comparisons between studies employing different methods should be avoided, at least until further evidence emerges in relation to the comparability of different methods. Moving forward, researchers should be more critical of the methods used to ensure the best or most accurate methods are used where possible.

References

1. White M. Food access and obesity. *Obes Rev* 2007;**8:** 99-107.

2. Wrigley N. 'Food deserts' in British cities: policy context and research priorities. *Urban Studies* 2002;**39:** 2029-2040.

3. Guy CM, David G. Measuring physical access to ‘healthy foods’ in areas of social deprivation: a case study in Cardiff. *International Journal of Consumer Studies* 2004;**28:** 222-234.

4. Walker RE, Keane CR, Burke JG. Disparities and access to healthy food in the United States: a review of food deserts literature. *Health & Place* 2010;**16:** 876-884.

5. Sparks AL, Bania N, Leete L. Comparative approaches to measuring food access in urban areas the case of Portland, Oregon. *Urban Studies* 2011;**48:** 1715-1737.

6. Dawson J, Marshall D, Taylor M, Cummins S, Sparks L, Anderson AS. Accessing healthy food: availability and price of a healthy food basket in Scotland. *Journal of Marketing Management* 2008;**24:** 893-913.

7. Black C, Moon G, Baird J. Dietary inequalities: What is the evidence for the effect of the neighbourhood food environment? *Health & Place* 2014;**27:** 229-242.

8. Maguire ER, Burgoine T, Monsivais P. Area deprivation and the food environment over time: a repeated cross-sectional study on takeaway outlet density and supermarket presence in Norfolk, UK, 1990-2008. *Health & Place* 2015;**33:** 142-147.

9. Burgoine T, Lake AA, Stamp E, Alvanides S, Mathers JC, Adamson AJ. Research report: Changing foodscapes 1980–2000, using the ASH30 Study. *Appetite* 2009;**53:** 157-165.

10. Jeffery RW, Baxter J, McGuire M, Linde J. Are fast food restaurants an environmental risk factor for obesity? *International Journal of Behavioral Nutrition and Physical Activity* 2006;**3:** 1-6.

11. World Obesity Federation. Trends in global obesity 2017 [cited 2018 01 May]. Available from: https://www.worldobesity.org/data/obesity-data-repository/resources/trends/12/.

12. Public Health England. Obesity and the environment: regulating the growth of fast food outlets 2014 [cited 2018 01 May]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/296248/Obesity_and_environment_March2014.pdf.

13. Local Government Association. Tipping the scales. Case studies on the use of planning powers to limit hot food takeaways. Local Government Association: London, 2016.

14. Greater London Authority. Takeaways toolkit 2012 [cited 2017 24 April]. Available from: https://www.london.gov.uk/sites/default/files/takeawaystoolkit.pdf.

15. Nykiforuk CIJ, Campbell EJ, Macridis S, McKennitt D, Atkey K, Raine KD. Adoption and diffusion of zoning bylaws banning fast food drive-through services across Canadian municipalities. *BMC Public Health* 2018;**18:** 1-12.

16. Diller PA, Graff S. Regulating food retail for obesity prevention: how far can cities go? *The Journal of Law, Medicine & Ethics* 2011;**39:** 89-93.

17. Waltham Forest Spatial Planning Unit. Hot Food Takeaway Supplementary Planning Document 2009 [cited 2016 27 May]. Available from: https://branding.walthamforest.gov.uk/documents/spd-hot-food-takeaway-mar10.pdf.

18. Cobb LK, Appel LJ, Franco M, Jones-Smith JC, Nur A, Anderson CAM. The relationship of the local food environment with obesity: a systematic review of methods, study quality, and results. *Obesity* 2015;**23:** 1331-1344.

19. Feng J, Glass TA, Curriero FC, Stewart WF, Schwartz BS. The built environment and obesity: a systematic review of the epidemiologic evidence. *Health & Place* 2010;**16:** 175-190.

20. Williams J, Scarborough P, Matthews A, Cowburn G, Foster C, Roberts N*, et al.* A systematic review of the influence of the retail food environment around schools on obesity-related outcomes. *Obes Rev* 2014;**15:** 359-374.

21. Gamba RJ, Schuchter J, Rutt C, Seto EY. 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:** 464-475.

22. Casey R, Oppert J-M, Weber C, Charreire H, Salze P, Badariotti D*, et al.* Determinants of childhood obesity: what can we learn from built environment studies? *Food Quality and Preference* 2014;**31:** 164-172.

23. Caspi CE, Sorensen G, Subramanian S, Kawachi I. The local food environment and diet: a systematic review. *Health & Place* 2012;**18:** 1172-1187.

24. Lytle LA, Sokol RL. Measures of the food environment: a systematic review of the field, 2007–2015. *Health & Place* 2017;**44:** 18-34.

25. Wilkins EL, Morris MA, Radley D, Griffiths C. Using Geographic Information Systems to measure retail food environments: discussion of methodological considerations and a proposed reporting checklist (Geo-FERN). *Health & Place* 2017;**44:** 110-117.

26. Charreire H, Casey R, Salze P, Simon C, Chaix B, Banos A*, et al.* Measuring the food environment using geographical information systems: a methodological review. *Public Health Nutr* 2010;**13:** 1773-1785.

27. Clary CM, Ramos Y, Shareck M, Kestens Y. 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:** 83-87.

28. Polsky JY, Moineddin R, Dunn JR, Glazier RH, Booth GL. Absolute and relative densities of fast-food versus other restaurants in relation to weight status: Does restaurant mix matter? *Prev Med* 2016;**82:** 28-34

29. Feng X, Astell-Burt T, Badland H, Mavoa S, Giles-Corti B. Modest ratios of fast food outlets to supermarkets and green grocers are associated with higher body mass index: Longitudinal analysis of a sample of 15,229 Australians aged 45 years and older in the Australian National Liveability Study. *Health & Place* 2018;**49:** 101-110.

30. Clary C, Lewis DJ, Flint E, Smith NR, Kestens Y, Cummins S. The local food environment and fruit and vegetable intake: A geographically weighted regression approach in the ORiEL study. *Am J Epidemiol* 2016;**184:** 837-846.

31. Mendez DD, Kim KH, Hardaway CR, Fabio A. Neighborhood racial and socioeconomic disparities in the food and alcohol environment: are there differences by commercial data sources? *Journal Of Racial And Ethnic Health Disparities* 2016;**3:** 108-116.

32. Hobbs M, Green M, Griffiths C, Jordan H, McKenna J. How different data sources and definitions of neighbourhood influence the association between food outlet availability and body mass index: a cross-sectional study. *Perspectives in Public Health* 2016;**137:** 158-161.

33. Thornton LE, Pearce JR, Macdonald L, Lamb KE, Ellaway A. Does the choice of neighbourhood supermarket access measure influence associations with individual-level fruit and vegetable consumption? A case study from Glasgow. *International Journal Of Health Geographics* 2012;**11:** 1-12.

34. Bernsdorf KA, Lau CJ, Andreasen AH, Toft U, Lykke M, Glümer C. Accessibility of fast food outlets is associated with fast food intake. A study in the Capital Region of Denmark. *Health & Place* 2017;**48:** 102-110.

35. Fiechtner L, Mona S, Sequist T, Block J, Duncan DT, Melly SJ*, et al.* Food environments and childhood weight status: Effects of neighborhood median income. *Childhood Obesity* 2015;**11:** 260-268.

36. Thomsen MR, Nayga RM, Alviola PA, Rouse HI. The effect of food deserts on the body mass index of elementary school children. *American Journal of Agricultural Economics* 2016;**98:** 1-18.

37. Burgoine T, Forouhi NG, Griffin SJ, Brage S, Wareham NJ, Monsivais P. Does neighborhood fast-food outlet exposure amplify inequalities in diet and obesity? A cross-sectional study. *Am J Clin Nutr* 2016;**103:** 1540-1547.

38. Reitzel LR, Regan SD, Nga N, Cromley EK, Strong LL, Wetter DW*, et al.* Density and proximity of fast food restaurants and body mass index among African Americans. *Am J Public Health* 2014;**104:** 110-116.

39. Wong MS, Chan KS, Jones-Smith JC, Colantuoni E, Thorpe RJ, Bleich SN. The neighborhood environment and obesity: Understanding variation by race/ethnicity. *Prev Med* 2017;**111:** 371-377.

40. Dwicaksono A, Brissette I, Birkhead GS, Bozlak CT, Martin EG. Evaluating the contribution of the built environment on obesity among New York State students. *Health Educ Behav* 2017;**45:** 1–12.

41. Popkin B, Adair L, Ng SW. Global nutrition transition and the pandemic of obesity in developing countries. *Nutr Rev* 2012;**70:** 3-21.

42. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. *Medicine & science in sports & exercise* 2000;**32:** 963-975.

43. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M*, et al.* The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised

studies in meta-analyses n.d. [cited 2018 6 Aug]. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

44. Li F, Harmer P, Cardinal BJ, Bosworth M, Johnson-Shelton D. Obesity and the built environment: does the density of neighborhood fast-food outlets matter? *Am J Health Promot* 2009;**23:** 203-209.

45. Zhao Z, Kaestner R, Xu X. Spatial mobility and environmental effects on obesity. *Econ Hum Biol* 2014;**14:** 128-140.

46. Gibson DM. The neighborhood food environment and adult weight status: estimates from longitudinal data. *Am J Public Health* 2011;**101:** 71-78.

47. An R, Sturm R. School and residential neighborhood food environment and diet among California youth. *Am J Prev Med* 2012;**42:** 129-135.

48. Lopez RP. Neighborhood risk factors for obesity. *Obesity (Silver Spring, Md)* 2007;**15:** 2111-2119.

49. Chen H-J, Wang Y. Changes in the neighborhood food store environment and children's body mass index at peripuberty in the United States. *J Adolesc Health* 2016;**58:** 111-118.

50. Shier V, An R, Sturm R. Is there a robust relationship between neighbourhood food environment and childhood obesity in the USA? *Public Health* 2012;**126:** 723-730.

51. Powell LM, Han E, Zenk SN, Khan T, Quinn CM, Gibbs KP*, et al.* Field validation of secondary commercial data sources on the retail food outlet environment in the US. *Health & place* 2011;**17:** 1122-1131.

52. Liese AD, Colabianchi N, Lamichhane AP, Barnes TL, Hibbert JD, Porter DE*, et al.* Validation of 3 food outlet databases: completeness and geospatial accuracy in rural and urban food environments. *Am J Epidemiol* 2010;**172:** 1324-1333.

53. Lake AA, Burgoine T, Greenhalgh F, Stamp E, Tyrrell R. The foodscape: classification and field validation of secondary data sources. *Health & Place* 2010;**16:** 666-673.

54. Burgoine T, Harrison F. Comparing the accuracy of two secondary food environment data sources in the UK across socio-economic and urban/rural divides. *International Journal of Health Geographics* 2013;**12:** 1-8.

55. Mason KE, Bentley RJ, Kavanagh AM. Fruit and vegetable purchasing and the relative density of healthy and unhealthy food stores – evidence from an Australian multilevel study. *J Epidemiol Community Health* 2013;**67:** 231-236.

56. Maguire ER, Burgoine T, Penney TL, Forouhi NG, Monsivais P. Does exposure to the food environment differ by socioeconomic position? Comparing area-based and person-centred metrics in the Fenland Study, UK. *International Journal of Health Geographics* 2017;**16:** 1-14.

57. Fan JX, Hanson HA, Zick CD, Brown BB, Kowaleski-Jones L, Smith KR. 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. *BMJ Open* 2014;**4**.

58. Burgoine T, Alvanides S, Lake AA. Creating 'obesogenic realities'; do our methodological choices make a difference when measuring the food environment? *International Journal of Health Geographics* 2013;**12:** 1-9.

59. Sterne JA, Smith GD. Sifting the evidence—what's wrong with significance tests? *Phys Ther* 2001;**81:** 1464-1469.

60. Bivoltsis A, Cervigni E, Trapp G, Knuiman M, Hooper P, Ambrosini GL. Food environments and dietary intakes among adults: does the type of spatial exposure measurement matter? A systematic review. *International Journal of Health Geographics* 2018;**17:** 1-20.

61. Thornton LE, Bentley RJ, Kavanagh AM. Fast food purchasing and access to fast food restaurants: a multilevel analysis of VicLANES. *International journal of behavioral nutrition and physical activity* 2009;**6:** 28.

62. Vogel C, Lewis D, Ntani G, Cummins S, Cooper C, Moon G*, et al.* The relationship between dietary quality and the local food environment differs according to level of educational attainment: A cross-sectional study. *PLoS One* 2017;**12:** 1-16.

63. Pearl J. Causal inference in statistics: An overview. *Statistics Surveys* 2009;**3:** 96-146.

64. James P, Berrigan D, Hart JE, Hipp JA, Hoehner CM, Kerr J*, et al.* Effects of buffer size and shape on associations between the built environment and energy balance. *Health & Place* 2014;**27:** 162-170.

65. Bodor JN, Rice JC, Farley TA, Swalm CM, Rose D. The association between obesity and urban food environments. *Journal Of Urban Health: Bulletin Of The New York Academy Of Medicine* 2010;**87:** 771-781.

66. Burgoine T, Forouhi NG, Griffin SJ, Wareham NJ, Monsivais P. Associations between exposure to takeaway food outlets, takeaway food consumption, and body weight in Cambridgeshire, UK: population based, cross sectional study. *BMJ* 2014;**348:** 1-10.

67. Christian WJ. Using geospatial technologies to explore activity-based retail food environments. *Spatial and Spatio-temporal Epidemiology* 2012;**3:** 287-295.

68. Crawford TW, Jilcott Pitts SB, McGuirt JT, Keyserling TC, Ammerman AS. Conceptualizing and comparing neighborhood and activity space measures for food environment research. *Health & Place* 2014;**30:** 215-225.

69. Lucan SC, Maroko AR, Patel AN, Gjonbalaj I, Abrams C, Rettig S*, et al.* Change in an urban food environment: storefront sources of food/drink increasing over time and not limited to food stores and restaurants. *Journal of the Academy of Nutrition and Dietetics*. Elsevier, 2018.

Tables

Table 1. General Study Characteristics

		Full Sample		Cobb Papers		New papers	
		N	%	N	%	N	%
Country							
	US	93	82.3	59	89.4	34	72.3
	Canada	10	8.8	7	10.6	3	6.4
	UK	5	4.4	-	-	5	10.6
	Australia	3	2.7	-	-	3	6.4
	Germany	1	0.9	-	-	1	2.1
	New Zealand	1	0.9	-	-	1	2.1
Gender							
	Mixed	100	88.5	59	89.4	41	87.2
	Females only	13	11.5	7	10.6	6	12.8
Age							
	Mixed	1	1.8	1	1.5	1	2.1
	Adults	75	66.4	45	68.2	31	63.8
	Children	36	31.9	20	30.3	16	34.0
Race							
	Mixed	73	64.5	43	65.2	30	63.8
	Mostly ethnic minority	10	8.8	4	6.1	6	12.8
	Mostly white	10	8.8	7	10.6	3	6.4
	Not reported	20	17.7	12	18.2	8	17.0
SES							
	Mixed	71	62.8	43	65.2	28	59.6
	High	2	1.8	1	1.5	1	2.1
	Low	19	16.8	11	16.7	8	17.0
	Not reported	21	18.6	11	16.7	10	21.3
Urban/rural							
	Mixed	33	29.2	20	30.3	13	17.7
	Urban	30	26.5	1	1.5	12	25.5
	Rural	3	2.7	18	27.3	2	4.3
	Not reported	47	41.6	27	40.9	20	42.6
Design							
	Cross-sectional	87	73.5	56	84.8	31	66.0
	Longitudinal	26	23.0	10	15.2	16	34.0

N = number of studies; SES = socioeconomic status; New papers = papers published after the Cobb review.

Table 2. Counts of area-based measures of the RFE

		UNIT OF MEASUREMENT									
		Count/area	**Count/ population**	**Non-standardised count**	**Presence vs absence**	**Relative**	**Audit score**	**Variety**	**Other**	**Unclear**	**TOTAL**
GEOGRAPHIC SCOPE	**Areal (N = 41)**										
	Census block group	3		1					1		**5**
	Census tract	4	5	8	4	3	1	2	3		**30**
	Zip code	1	3	1					2		**7**
	Researcher defined	4	3	1	1	2		1			**12**
	Other - administrative	1							1	1	**2**
	Buffer – Euclidian (N = 23)										
	<=200m	3*	1	n/a	1						**5**
	>200-400m	5*	1	n/a	1						**7**
	>400-800m	14*	2	n/a	2	1					**19**
	>800-1600m	15*	2	n/a	4	3	1	1	1	3	**30**
	>1600-2400m	5*	1	n/a							**6**
	>2400-3200m	5*	1	n/a							**6**
	>3200m	5*	1	n/a	1				1	1	**9**
	Buffer – Network (N = 27)										
	<=200m										**0**
	>200-400m	1	1	2	3						**5**
	>400-800m	3		8	3	2	3		1		**12**
	>800-1600m	2	1	4	5	4	4	1	2		**19**
	>1600-2400m			3			2				**2**
	>2400-3200m	1		3	2	1			1		**5**
	>3200m			5	3		3		1		**7**
	Buffer – Undefined (N = 8)										
	>200-400m	1		1	1				1		**4**
	>400-800m			1							**1**
	>800-1600m			2	1	1			2		**5**
	>1600-2400m	1		4	2	1			2		**10**
	>2400-3200m			1		1					**2**
	>3200m			2	1	1			2		**5**
	TOTAL	**74**	**22**	**46**	**35**	**20**	**14**	**5**	**21**	**5**	

Note. Many studies employed multiple measures of the RFE, and thus the total number of measures (242) exceeds the number of studies (113). N = number of studies employing each broad method. Non-standardised count = measures of the raw counts of outlets that are not standardised e.g. to a given area or population. Relative = measures of the availability of one outlet type relative to one or more other outlet types. Audit score = measures derived from *within-store audits e.g. the total shelf space devoted to fruits and vegetables within a buffer. Variety = measures of the number of different outlet types. Other = other measures of the RFE, including counts of outlets relative to the length of roads within a buffer, weighted counts of outlets and counts per area per population. Buffer – Undefined = studies that described using a buffer measure of the RFE, but did not describe whether this was a network or Euclidian buffer.*

**Measures of the raw count of outlets within Euclidian buffers were classified as count/area, because Euclidian buffers of a given radius have a fixed area.*

Table 3. GeoFERN reporting quality by domain.

GeoFERN domains & associated marking criteria	Full details N (%)	Partial details N (%)	No details N (%)	Not applicable
DOMAIN 1: DATA SOURCE				
Overall	55 (48.7%)	47 (41.6%)	11 (9.7%)	-
Name of data creator reported?	79 (76.0%)	5 (4.8%)	20 (19.2%)	9[(1)]
Dataset name reported?	42 (58.3%)	6 (8.3%)	24 (33.3%)	4[(1)]
Publication date reported[(2)]	71 (62.8%)	3 (2.7%)	39 (34.5%)	0
DOMAIN 2: EXTRACTION METHODS				
Overall	28 (27.5%)	31 (30.4%)	43 (42.1%)	11[(3)]
Extraction Method Reported	54 (52.9%)	11 (10.8%)	37 (36.3%)	11[(3)]
Search Terms Reported	53 (52.0%)	17 (16.7%)	32 (31.4%)	11[(3)]
DOMAIN 3: CONSTRUCT DEFINITIONS				
Overall	18 (15.9)	95 (84.1)	0	-
Construct names listed	112 (99.1)	1 (0.9)	0	-
Construct scope clear	60 (53.1)	26 (23.0)	27 (23.9)	-
Classification method clear	51 (45.1)	25 (22.1)	37 (32.7)	-
Examples provided	37 (32.7)	13 (11.5)	63 (55.8)	-
DOMAIN 4: GEOCODING METHODS				
Overall	3 (2.7)	103 (91.2)	7 (6.2)	-
Clear whether geocoding used?	61 (54.0)	32 (28.3)	20 (17.7)	-
Address model reported?	15 (13.3)	28 (24.8)	63 (55.8)	7[(4)]
Match rate reported	83 (73.5)	-	23 (20.4)	7[(4)]
Software reported	48 (42.5)	1 (0.9)	57 (50.4)	7[(4)]
Urban/rural reported	22 (19.5)	44 (38.9)	47 (41.6)	-
DOMAIN 5: RFE METRICS				
Overall	22 (19.5)	91 (80.5)	0	-
Conceptual environment defined	110 (97.3)	-	3 (2.7)	-
Areal (N = 41)				
Type of zone defined?	40 (97.6)	-	1 (2.4)	-
Boundary data reported?	14 (37.8)	-	23 (62.2)	4[(5)]
Intensity metric reported?	38 (92.7)	-	3 (7.3)	-
Buffer (N = 60)				
Buffer type defined?	52 (86.7)	-	8 (13.3)	-
Buffer size defined?	60 (100.0)	-	0	-
Intensity metric defined?	59 (98.3)	-	1 (1.7)	-
Proximity (N = 37)				
Proximity type defined?	32 (86.5)	-	5 (13.5)	-
Network (N = 46)				
Types of roads/paths described	43 (93.5)	-	3 (6.5)	-
Network data cited?	7 (15.2)	-	39 (84.8)	-
Gravity (N = 5)				
Radius defined?	4 (80.0)	-	1 (20.0)	-
Decay function defined?	4 (80.0)	-	1 (20.0)	-
Other (N = 5)				
Metric clearly described?	5 (100.0)	-	0	-

Table shows the number (N) of studies that reported full, partial or none of the details listed as essential within the GeoFERN reporting checklist (the 'marking criteria') for each GeoFERN domain (23) ('Overall' score). Also shown is a break-down of the reporting quality (i.e. numbers of studies providing full, partial or no details) for each specific marking criterion. The 'data creator' is the person or entity that created the RFE data e.g. 'Dun & Bradstreet'. 'Extraction methods' refer to the methods used to extract food outlets of interest from a wider dataset which may contain food outlets not of interest and/or non-food businesses. Often search terms are used, which might include outlet names or classifications within the dataset. Outlet 'constructs' include e.g. 'fast food *outlets', 'supermarkets' and 'convenience stores'. RFE = Retail Food Environment. Percentages shown are the percentages of eligible studies for each marking criterion; excluding those for which GeoFERN marking criterion was not applicable.*

(1) Not applicable because food environment data was collected through street audits, or because dataset does not have a name.

(2) Or date of audit reported if data was collected via street audit.

(3) Not applicable because data was collected via street audits, or no data extraction was required.

(4) Not applicable because geocoding was not required.

(5) Boundary data was not applicable for 4 studies; wherein areal units were researcher-defined.

Table 4. Numbers of statistically significant positive, negative and null associations between RFE and weight status.

Exposure-population grouping	**Positive***		**Null**		**Negative***		**Total**	**No. of**
	N	%	**N**	%	**N**	%	**N**	**studies**
FAST FOOD OUTLETS								
Full sample	84	20.8%	303	75.0%	17	4.2%	404	74
High quality studies only	12	30.8%	26	66.7%	1	2.6%	39	8
Adults	52	22.1%	173	73.6%	10	4.3%	235	44
- Males	3	9.4%	24	75.0%	5	15.6%	32	10
- Females	13	25.5%	34	66.7%	4	7.8%	51	16
- Low SES	4	22.2%	14	77.8%	0	0.0%	18	6
- High SES	1	7.7%	12	92.3%	0	0.0%	13	5
- White	2	9.1%	19	86.4%	1	4.5%	22	7
- Urban	13	29.5%	31	70.5%	0	0.0%	44	13
Children	32	19.5%	126	76.8%	6	3.7%	164	28
- Low SES	12	37.5%	19	59.4%	1	3.1%	32	10
- Non-white	0	0.0%	8	88.9%	1	11.1%	9	5
- Urban	2	4.7%	39	90.7%	2	4.7%	43	9
CONVENIENCE STORES								
Full sample	30	10.9%	218	79.6%	26	9.5%	274	52
High quality studies only	0	0.0%	47	95.9%	2	4.1%	49	7
Adults	11	6.9%	125	78.6%	23	14.5%	159	31
- Males	1	3.4%	20	69.0%	8	27.6%	29	8
- Females	7	12.7%	43	78.2%	5	9.1%	55	14
- Low SES	1	5.0%	19	95.0%	0	0.0%	20	7
- White	2	8.3%	19	79.2%	3	12.5%	24	5
- Urban	2	5.7%	26	74.3%	7	20.0%	35	11
- Rural	0	0.0%	18	100.0%	0	0.0%	18	5
Children	19	16.5%	93	80.9%	3	2.6%	115	21
- Females	4	16.0%	21	84.0%	0	0.0%	25	5
- Low SES	11	39.3%	16	57.1%	1	3.6%	28	8
- Non-white	7	33.3%	12	57.1%	2	9.5%	21	6
- Urban	2	10.5%	16	84.2%	1	5.3%	19	5
SUPERMARKETS/GROCERY STORES								
Full sample	37	6.6%	454	80.5%	73	12.9%	564	70
High quality studies only	5	4.0%	108	85.7%	13	10.3%	126	12
Adults	34	9.0%	293	77.7%	50	13.3%	377	46
- Males	0	0.0%	36	83.7%	7	16.3%	43	9
- Females	17	12.6%	102	75.6%	16	11.9%	135	20
- Low SES	5	7.4%	60	88.2%	3	4.4%	68	12
- White	7	12.3%	47	82.5%	3	5.3%	57	5
- Urban	6	6.3%	74	77.9%	15	15.8%	95	16
- Rural	7	15.9%	37	84.1%	0	0.0%	44	5
Children	3	1.6%	157	86.3%	22	12.1%	182	23
- Low SES	0	0.0%	17	89.5%	2	10.5%	19	8
- Urban	0	0.0%	24	92.3%	2	7.7%	26	8
RESTAURANTS								
Full sample	11	6.3%	131	75.3%	32	18.4%	174	29
High quality studies only	4	9.8%	34	82.9%	3	7.3%	41	5
Adults	1	1.0%	71	70.3%	29	28.7%	101	18
- Males	0	0.0%	18	62.1%	11	37.9%	29	8
- Females	0	0.0%	24	72.7%	9	27.3%	33	11
- Urban	0	0.0%	16	76.2%	5	23.8%	21	5

Children	10	13.7%	60	82.2%	3	4.1%	73	11

*N = number of associations. SES = socioeconomic status. *'Positive associations' refer to statistically significant ($p < 0.05$) associations indicating increased access/exposure to food outlets is associated with increased obesity, and 'negative associations' refer to statistically significant associations indicating increased access/exposure to food outlets is associated with decreased obesity.*

Table 5. Numbers of statistically significant positive, negative and null associations between RFE and weight status, stratified by measurement method

Exposure-method grouping	Positive*		Null		Negative*		Total	No. of
	N	%	N	%	N	%	N	studies
DATA SOURCE								
Fast Food Outlets								
- Government	13	16.0%	65	80.2%	3	3.7%	81	19
- Commercial	59	24.5%	170	70.5%	12	5.0%	241	36
Supermarket/Grocery								
- Government	16	8.8%	148	81.3%	18	9.9%	182	20
- Commercial	5	1.8%	229	81.5%	47	16.7%	281	28
EXTRACTION METHODS								
Fast Food Outlets								
- Proprietary classifications	54	26.3%	142	69.3%	9	4.4%	205	30
- Combination of methods	8	16.3%	39	79.6%	2	4.1%	49	13
Supermarket/Grocery								
- Proprietary classifications	16	4.8%	272	82.2%	43	13.0%	331	24
- Combination of methods	5	6.3%	56	70.9%	18	22.8%	79	14
OUTLET CLASSIFICATIONS								
Fast Food Outlets								
- All	84	20.8%	303	75.0%	17	4.2%	404	74
- Narrow	24	26.1%	66	71.7%	2	2.2%	92	12
- Moderate	29	22.3%	97	74.6%	4	3.1%	130	26
- Broad	12	19.4%	48	77.4%	2	3.2%	62	10
Supermarket**								
- All	37	6.6%	454	80.5%	73	12.9%	564	70
- Narrow	2	1.7%	90	77.6%	24	20.7%	116	18
- Moderate	2	2.9%	59	86.8%	7	10.3%	68	14
- Broad	7	10.4%	54	80.6%	6	9.0%	67	9
NEIGHBOURHOOD DEFINITION								
Fast Food Outlets - Adults								
- Areal measures	10	17.9%	45	80.4%	1	1.8%	56	18
- Person-centric measures	42	23.5%	128	71.5%	9	5.0%	179	31
- Buffer <1km	9	36.0%	16	64.0%	0	0.0%	25	6
- Buffer 1-2km	10	18.5%	43	79.6%	1	1.9%	54	17
- Buffer > 2km	15	34.1%	28	63.6%	1	2.3%	44	11
Fast Food Outlets - Children								
- Areal measures	6	18.8%	24	75.0%	2	6.3%	32	7
- Person-centric measures	26	19.7%	102	77.3%	4	3.0%	132	23
- Buffer ≤ 400m	10	25.0%	30	75.0%	0	0.0%	40	8
- Buffer > 400m	1	2.0%	45	91.8%	3	6.1%	49	13
Supermarket/Grocery - Adults								
- Areal measures	11	12.2%	71	78.9%	8	8.9%	90	19
- Person-centric measures	23	8.0%	222	77.4%	42	14.6%	287	33
- Buffer <1km	3	8.3%	30	83.3%	3	8.3%	36	6
- Buffer 1-2km	8	10.8%	53	71.6%	13	17.6%	74	15
- Buffer > 2km	6	7.3%	62	75.6%	14	17.1%	82	8
Supermarket/Grocery - Children								
- Areal measures	2	2.9%	62	88.6%	6	8.6%	70	6
- Person-centric measures	1	0.9%	95	84.8%	16	14.3%	112	19
- Buffer <=400m	0	0.0%	17	100.0%	0	0.0%	17	5
- Buffer > 400m	1	1.9%	49	94.2%	2	3.8%	52	9
METRIC TYPE								
Fast Food Outlets								
- Count (non-standardised)	7	18.9%	25	67.6%	5	13.5%	37	17
- Count/area	31	22.8%	104	76.5%	1	0.7%	136	26
- Count/capita	7	20.0%	28	80.0%	0	0.0%	35	11
- Presence/absence	0	0.0%	25	96.2%	1	3.8%	26	8
- Proximity	22	28.6%	53	68.8%	2	2.6%	77	25
Supermarket/Grocery								
- Count (non-standardised)	8	6.6%	102	83.6%	12	9.8%	122	16
- Count/area	9	5.7%	131	83.4%	17	10.8%	157	17
- Count/capita	4	7.5%	40	75.5%	9	17.0%	53	10
- Presence/absence	6	18.8%	23	71.9%	3	9.4%	32	11
- Proximity	6	6.3%	75	78.1%	15	15.6%	96	26

ALTERNATIVE MEASURES								
Healthy – composite	2	3.0%	48	72.7%	16	24.2%	66	9
Unhealthy – relative	23	21.3%	85	78.7%	0	0.0%	108	12
Healthy – score	9	9.2%	74	75.5%	15	15.3%	98	7
Total outlets	3	7.7%	33	84.6%	3	7.7%	39	8
Total restaurants (including fast food)	2	8.3%	19	79.2%	3	12.5%	24	8
Total food stores	3	8.6%	30	85.7%	2	5.7%	35	5

N = number of associations. Supplement 6 provides details on definitions of 'narrow', 'moderate' and 'broad' scope.

**'Positive associations' refer to statistically significant (p < 0.05) associations indicating increased access/exposure to food outlets is associated with increased obesity, and 'negative associations' refer to statistically significant associations indicating increased access/exposure to food outlets is associated with decreased obesity.*

***Excludes grocery stores, unless these were included under the same classification as supermarkets.*

Figures

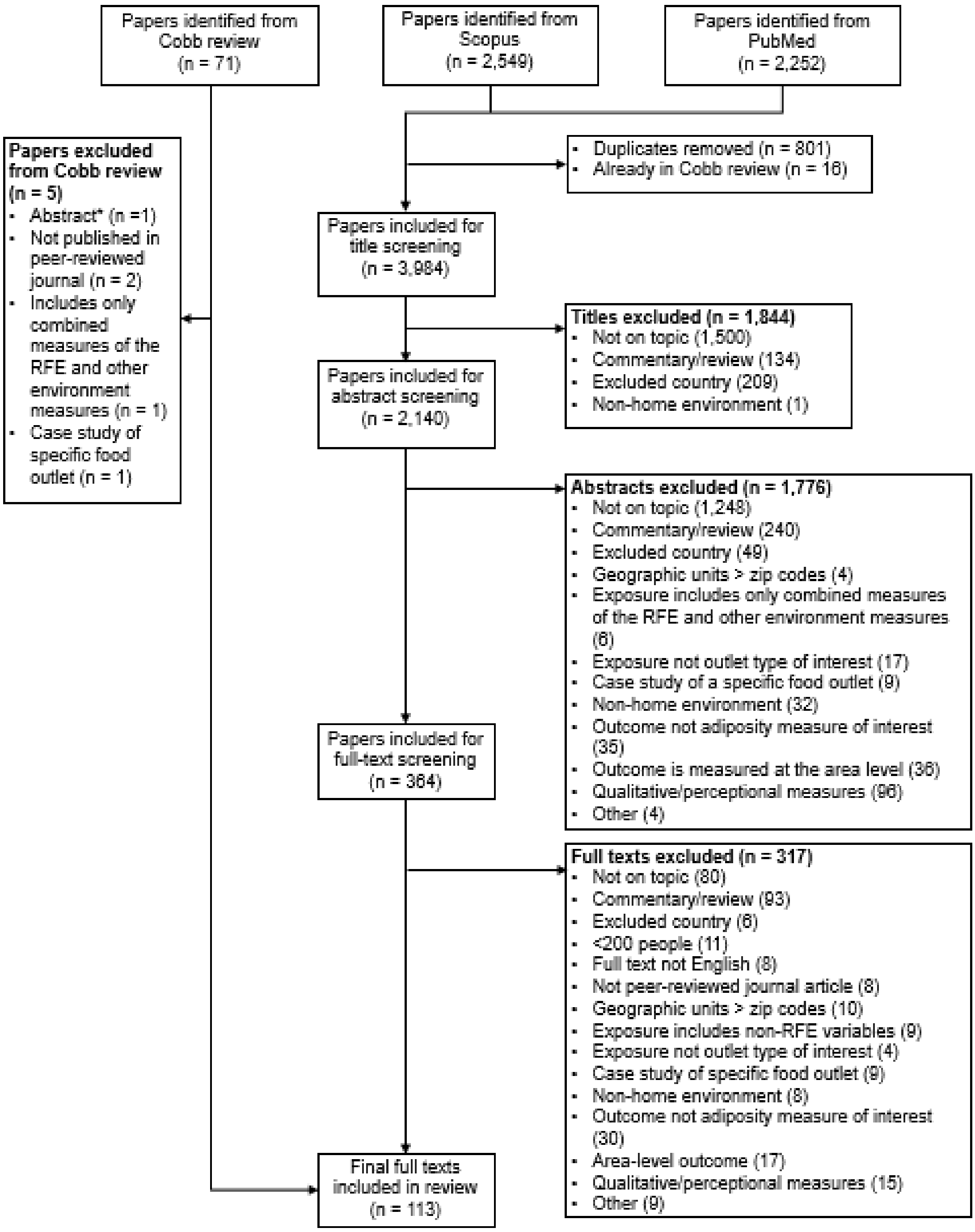

*Figure 1. Flow chart illustrating screening process for this review. RFE = Retail Food Environment. Note, for the papers excluded from the Cobb review, the third and fourth criteria listed above were also applied in the original Cobb review, but appeared to have been incorrectly applied in respect of two papers. *Article was an abstract corresponding to a full-text paper identified in the top-up search and thus was excluded to avoid duplication.*

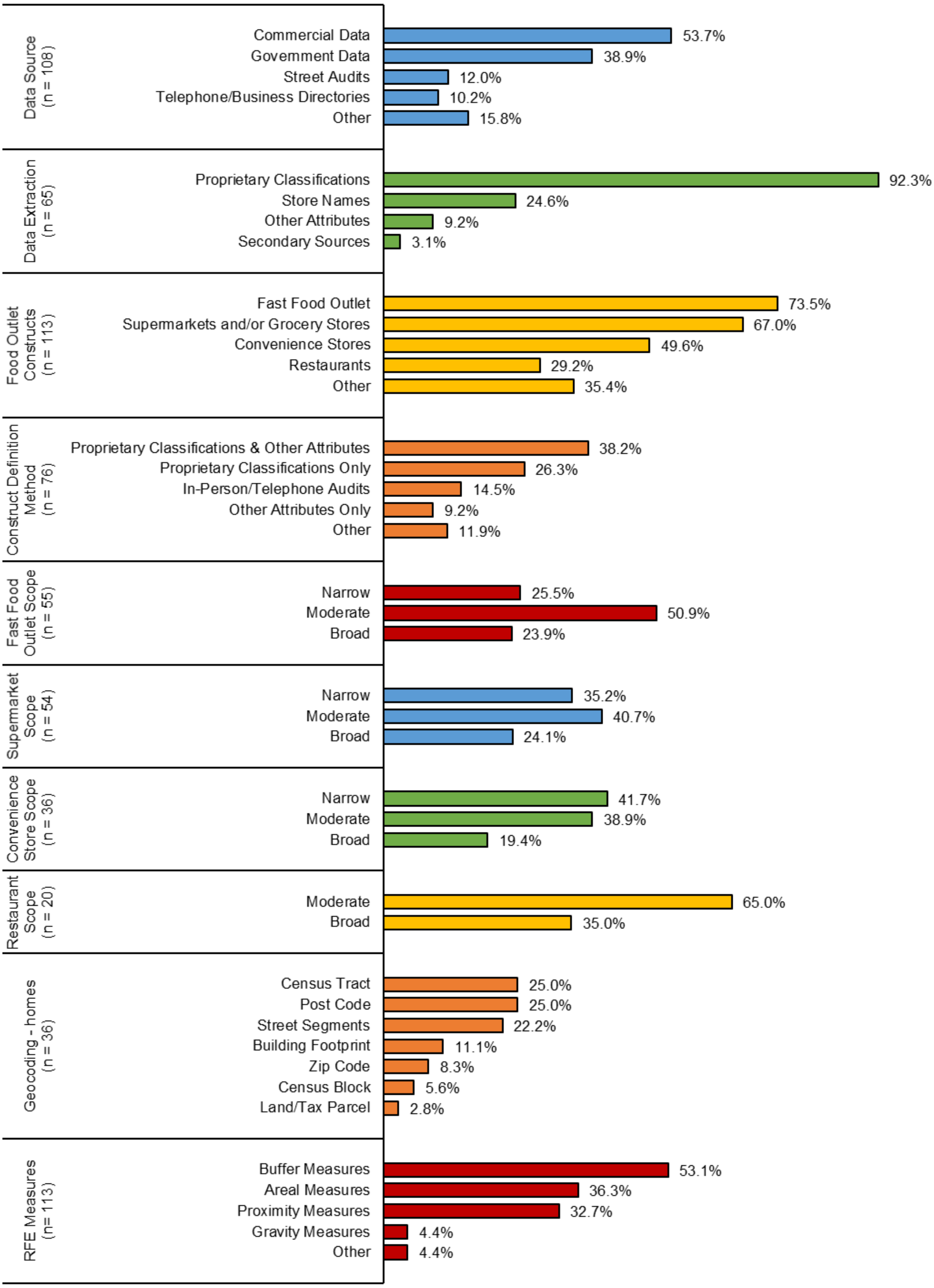

Figure 2. Methods used within studies. n = number of studies for which methodological aspect is applicable and for which sufficient methodological information was provided. RFE = Retail Food Environment. Some studies used more than one method within a given methodological aspect, and thus percentages shown do not always add up to 100%. 'Other' data sources included internet

searching, data from national mapping agencies, and satellite imagery. 'Other' food outlet constructs included various composite measures such as supermarkets and greengrocers combined, or fast food outlets and convenience stores combined. 'Other attributes' used for construct definition was limited to information contained within the RFE dataset and included outlet name, size, number of employees or tills. 'Other' methods for applying outlet constructs included use of supplementary information e.g. websites, and interviews with local residents. 'Other' RFE metrics included e.g. a binary measure of whether the neighbourhood centroid was closer to a supermarket or ethnic market and measures of relative store 'attractiveness' (accounting for distance and store size).

Supplementary Materials – Supplements 1-7 & 10

Supplement 1. Search Strategy

Pubmed search

Search terms:

(“Overnutrition”[Mesh] OR “Overweight”[Mesh] OR overnutrition[tw] OR “over nutrition”[tw] OR overweight[tw] OR “over weight”[tw] OR obes*[tw] OR “Body Mass Index”[Mesh] OR “Body Mass Index”[tw] OR BMI[tw] or “waist circumference”[tw] or weight change*[tw])

AND

(“Food Industry”[Mesh:NoExp] OR “Food Supply”[Mesh] OR “Food Services”[Mesh] OR “Environment Design”[Mesh] OR grocer*[tw] OR supermarket*[tw] OR super market*[tw] OR food store*[tw] OR corner store*[tw] OR convenience store*[tw] OR food environment*[tw] OR “Residence Characteristics”[Mesh] OR food outlet*[tw] OR “fast food”[tw] OR restaurant*[tw] OR carryout*[tw] OR takeaway*[tw] OR “food supply”[tw] OR food desert*[tw] OR food swamp*[tw] OR “food and physical activity environment”[tw] OR “food availability”[tw] OR “food access”[tw] OR built environment*[tw] OR “food and physical activity environments”[tw])

AND

("2014/01/01"[Date - Create] : "3000"[Date - Create])

Results were filtered to include papers published after 01/01/90 only – in line with search dates of the original Cobb review.

Scopus search

Search for the following in title, abstract or key words:

(obes* OR overweight OR “over weight” OR BMI OR “body mass index” OR “waist circumference” OR “weight change” OR “weight gain” OR “weight loss”)

AND

(grocer* OR supermarket* OR “food store*” OR “corner store*” OR “convenience store*” OR “food environment” OR “food outlet* OR “fast food” OR restaurant* OR carryout OR takeaway OR “food desert” OR “food swamp” OR “food access” OR “food availability” OR “food and physical activity environment*”)

Results were filtered to include only those papers published during or after 2013 (there is no option in Scopus to search by date added, and it is also not possible to limit inclusion date to a specific month).

Supplement 2. Data Extraction

Data were extracted for the following data fields (items in bold indicate data extracted for the purpose of assigning a GeoFERN reporting score):

Study Design Characteristics

- Sample size (free-text)
- Country of origin (free-text)
- Study name (free-text)
- Population demographics (free-text)
- Age classification (drop-down options)
- Gender classification (drop-down options)
- Race classification (drop-down options)
- Socio-economic classification (drop-down options)
- Longitudinal/cross-sectional design (drop-down options)

Data Source (GeoFERN Domain 1)

- Exposure data source type (drop-down options)
- Data source details (free text)
- Uses Dun & Bradstreet or InfoUSA (yes/no)
- Uses business patterns data or local authority licensing/hygiene data (yes/no)
- **Name of data creator reported (yes/no)**
- **Name of dataset reported (yes/no)**
- **Publication date of dataset reported (yes/no)**

Data Extraction (GeoFERN Domain 2)

- **Food outlet extraction method reported (yes/no)**
- Food outlet data extraction method (drop-down options)
- Food outlet data extraction method details e.g. specific search terms used (free text)
- **Search terms listed (where applicable) (yes/no)**
- List of search terms used (free text)

Food Outlet Constructs (GeoFERN Domain 3)

- Food outlet construct names (free text)
- Fast food outlets investigated (yes/no)
- Supermarkets/grocery stores investigated (yes/no)
- Restaurants investigated (yes/no)
- Convenience stores investigated (yes/no)
- Composite measures investigated (yes/no)
- Unknown/undefined outlet types investigated (yes/no)
- Food outlet construct definition method (drop-down options)
- Food outlet construct definition details e.g. scope and search terms (free text)
- Cites use of NAICS (North American Industry Classification Scheme) or SIC (Standard Industry Classification) scheme (yes/no plus details)
- Fast food outlet definition (free text)
- Fast food outlet scope (drop-down text)
- Supermarket definition (free text)
- Supermarket scope (drop-down text)

- Convenience store definition (free text)
- Convenience store scope (drop-down text)
- Restaurant definition (free text)
- Restaurant scope (drop-down text)
- **Construct names listed (yes/no)**
- **Scope of classifications adequately described (yes/no)**
- **Method used to apply classifications described (yes/no)**
- **Examples of outlets falling within/outside construct definitions provided (yes/no)**

Geocoding (GeoFERN Domain 4)

- **Was it clear whether geocoding was used? (yes/no)**
- **Was the address model reported for homes and food outlets were applicable (yes/no)**
- Geocoding methods used for homes (free-text)
- Geocoding methods used for food outlets (free-text)
- **Match rate reported? (yes/no)**
- **Geocoding software reported? (yes/no)**
- **Environmental context reported? (yes/no)**
- Urban/rural environmental context (drop-down options)

Access Metrics (GeoFERN Domain 5)

- Summary of measures used (free-text)
- **Conceptual environment clearly defined? (yes/no)**
- Areal Metrics
    - **Type of areal zoning system defined? (yes/no)**
    - Zone category (drop-down options)
    - Zone name (free text)
    - **Boundary data reported? (yes/no)**
    - **Intensity metric defined? (yes/no)**
    - Intensity metric type (drop-down options)
    - Intensity metric details (free text)
- Buffer Metrics
    - **Buffer type defined? (yes/no)**
    - Buffer type (drop-down options)
    - **Buffer size defined? (yes/no)**
    - Buffer size (free text)
    - **Intensity metric defined? (yes/no)**
    - Intensity metric type (drop-down options)
    - Intensity metric details (free text)
- Proximity Metric
    - **Distance type defined? (yes/no)**
    - Distance type (drop-down options)
    - Distance type details (free text)
- Network Metrics
    - **States types of paths/roads included? (yes/no)**
    - Types of paths/roads included (free text)
    - **Cites the network data used? (yes/no)**

    - Network data used (free text)
- Gravity Metrics
    - **Zone radius defined? (yes/no)**
    - Zone radius (free text)
    - **Decay function defined? (yes/no)**
    - Decay function (free text)
    - Type of intensity measure (free text)
    - Intensity measure details (free text)
- Other
    - **Metric clearly described? (yes/no)**
    - Other metric details (free text)

Outcomes

- Exposure (free text)
- Exposure type (drop-down option)
- Metric type (free text)
- Outcome (free text)
- Population (free text)
- No. of positive associations (number)
- No. of null associations (number)
- No. of negative associations (number)

Supplement 3. Defining the 'Main Model'

Often papers will describe and report results for multiple models. However, for this review, data was only extracted for the 'main model'. The following decision rules were applied in the order shown. In each case, care was taken to ensure results were extracted for models that aimed to understand associations between the RFE and obesity in preference to alternative research questions where the RFE might have been included in models e.g. as a covariate.

1. The main model is the model for which data is presented in a results table.
2. If data is presented for multiple models:
    a. In the case of multivariate vs bivariate regressions, the main model is the multivariate model.
    b. If there are several multivariate regressions, the main model is the most fully-adjusted model included in the main text (i.e. excluding models that are included in supplementary materials).
    c. If multiple models meet the above criteria, then the main model is the most complex model (e.g. a model that accounts for spatial dependence would be selected over one that does not). As an exception to this, models that include interaction terms with food environment measures were excluded, because the coefficients of the interaction terms could not be usefully be interpreted in this review.
    d. If multiple models are equally complex, data were extracted for both models.

Note, that models that were run for different access measures (e.g. buffer sizes/intensity metrics) and different sub-groups of people (e.g. low/high socioeconomic status) were considered to answer different research questions, and thus all such models were extracted.

Supplement 4. Data extraction agreement

Detailed information on the evaluation of agreement between reviewers in systematic reviews is rarely reported. It is unclear, for example, whether reviews generally award marks for agreement on a paper-by-paper basis or by a data field-by-data field basis. In the former case, data extracted would be required to agree across all data fields for a mark to be awarded. There is also no 'gold standard' or best practice method, although the Cochrane Handbook for systematic reviews does advise that "if agreement is assessed, this should be done only for the most important data (e.g. key risk of bias assessments, or availability of key outcomes)", and suggests that the Kappa statistic can be used for coded data. Data extraction was therefore assessed as set out below. We decided to mark agreement at the level of data fields, firstly because this is what the recommendations in the Cochrane Handbook seem to infer, and secondly because it was felt this gave a fairer and more detailed indication of the level of agreement between reviewers.

To assess agreement between data extracted on the study design and the methods sections of this review, the data extracted by the first reviewer was integrated in a spreadsheet together with the data extracted by the second reviewer (MH). Areas of disagreement were resolved through discussion. For each data field in the spreadsheet, a mark of '1' was awarded if the final agreed decision was in line with the initial decision made by the first reviewer; and a mark of 0 was awarded if the final agreed decision was different to the initial decision of the first reviewer. For example, if the first reviewer identified a sample size of 40,000, whereas the second reviewer identified a sample size of 39,567, then a mark of 1 would be awarded if the final agreed sample size was 40,000 whereas a mark of 0 would be awarded if the final agreed sample was anything other than 40,000 (irrespective of whether the final agreed sample was in line with that of the second reviewer). Agreement was reported as percentage agreement with the first reviewer to give an indication of the reliability of data extraction for studies that had not been double-screened. The Kappa Statistic was not used because some data fields were free-text responses and thus the Kappa statistic could not be used for these.

Agreement for the outcomes data (numbers of null and statistically significant associations) was assessed at the level of individual associations rather than papers, because the number of tested associations differed across studies. Areas of disagreement were identified, and resolved through discussion. Again, percentage agreement with the first reviewer's initial decision was deemed to be of most interest as indicating the reliability of the singly-screened papers. Thus, for each association identified by the first reviewer or agreed after discussion, 5 marks were available for agreement on (i) the exposure name (e.g. 'supermarket'), (ii) the type of exposure measure (e.g. count within census tract), (iii) the outcome (e.g. 'BMI'), (iv) the population (e.g. women) and (v) the direction of the association (positive, negative or null). A mark of 1 was awarded for each domain if the first reviewer's initial decision was in agreement with the final decision, and a mark of 0 was awarded if there was disagreement between the first author and the final decision. Thus, for example, an association which was identified by the first author, but was later determined to be erroneous would be awarded 0 out of 5 marks. An association that was erroneously not identified by the first author would also receive 0 out of 5 marks. An association identified by the second reviewer, but not the first and later determined to be erroneous was not included in the marking, because agreement with the second reviewer was not of primary interest.

Supplement 5. Risk of bias tool

Studies were appraised against 10 quality criteria, with a mark of 1 being awarded for each applicable quality criterion met. The quality criteria were:

1. Food environment data was collected through street audits, or if collected through secondary sources, data was validated via in-person audits or phone calls to outlets.
2. Study design controlled for neighbourhood self-selection bias.
3. RFE exposure was centred on the participant address (or UK postcode).
4. Height and weight were objectively measured.
5. Analyses control for age, race, sex and socio-economic status (or uses a study design not requiring this)
6. Controls for at least one of: neighbourhood socio-economic status or racial composition.
7. Employs a multi-level model or alternative method accounting for clustered standard errors when food environment exposures are at the area level rather than individual level (coded as n/a where individual-level food environment exposures are employed).
8. Does not control for variables on the causal pathway (such as food purchasing or consumption behaviours).
9. Data is presented for all main analyses describes in the statistical methods section.
10. Analyses are based on an explicit causal framework.

No mark was awarded for a criterion if it was unclear/not reported whether that criterion was met.

Quality scores were expressed as a percentage of the maximum available marks to account for some criteria being not applicable for some studies. Descriptive statistics are shown in Table S5 below.

Table S5. Percentage of studies meeting quality criteria

Quality Criteria	% studies meeting criterion
1. FO validated in person/on the phone	21.2%
2. Exposure centred on participant addresses	63.7%
3. Height and weight objectively measured	54.0%
4. Account for neighbourhood self-selection	14.2%
5. Controls for age, race, sex, SES	59.3%
6. Controls for at least one of neighbourhood SES/racial composition	61.1%
7. A multi-level model/alternative used where necessary	81.1%*
8. Avoids controlling for variables on causal pathway	89.4%
9. Data presented for all main analyses	67.3%
10. Uses causal framework	2.7%

No. of studies appraised = 113.
*of 53 studies for which this criterion was applicable.

Supplement 6. Data Synthesis

Data was coded as follows:

Demographics

- Gender: 'Women' (if >95% of sample is women), 'Mixed' (any other gender composition – note there were no studies that were >95% males).
- Age: 'Adults' (whole sample ≥18yrs), 'Children' (whole sample ≤18yrs), 'Mixed' (both below and over 18yrs)
- Race/ethnicity: 'Mostly ethnic minorities' (if ≥85% of sample are non-white Caucasian), 'Mostly white' (if ≥85% of sample are white Caucasian), Mixed (if <85% of sample is white Caucasian)
- Socioeconomic status: It was not possible to apply an objective definition of high/low SES, because measures of SES (e.g. education and income) were varied, and reported inconsistently across papers. Therefore, populations were classified as 'high SES' or 'low SES' only if the paper described population as high/low SES and the data appeared to support this OR if recruitment was such as to ensure participants came from a low-income neighbourhood OR were enrolled on a program indicating low SES (e.g. WIC participants). If SES data was reported, but the population did not meet the above requirements for high/low SES then the population was classified as 'mixed'. If no SES information was given, then SES was classified as 'not reported'.

Construct definition scope

Based on the extracted free-text data, the scope of the three most commonly investigated outlet types (supermarkets, convenience stores and fast food outlets) were classified as narrow, broad or moderate as set out below.

- Supermarkets. These classifications are designed to delineate between a broad definition of supermarkets which encompasses both large chain supermarkets, and grocery stores of all sizes, a moderate scope, which encompasses large chain supermarkets and mid-large sized grocery stores and narrow scopes which encompass only large chain supermarkets. Specific cut points e.g. for annual sales were selected to allow delineation between those commonly used in the literature. For example, requiring > 50 employees and/or >$5 million in sales was a very common criterion for defining large supermarkets.
    - Narrow: requires at least one of:
        - At least 50 employees
        - At least $5 million annual sales
        - To be a large chain/franchised supermarket or hypermarket
    - Moderate: typically retail a range of staple foods, including milk, bread, fresh and frozen fruit and vegetables, and include large/mid-sized grocery stores. May be required to have at least one of:
        - At least 15,000 feet in size
        - At least 4 cash registers
        - At least $1 million annual sales
        - At least 15 employees
    - Broad: Outlets classified with North American Industry Classification (NAICS) code 445110 ('supermarkets and other grocery stores excluding convenience stores'), or falling under a similar scope. May possibly have limited restrictions on size, turnover, employees etc. which fall below those required for 'moderate' (e.g. lower annual sales threshold).

- Fast food outlets. The cut-points for establishing whether the scope was narrow, moderate or broad were determined after appraising the range of scopes employed, with ‘moderate’ being the most commonly used scope, and narrow/broad being designed to capture the extremes.
    - Moderate: restaurants offering limited wait service and payment before food is served, broadly in line with the definition of ‘limited service restaurants’ according to the NAICS scheme (classification code 722211). May additionally include pizzerias. Outlet may offer foods for takeaway only (e.g. kebab shop, fish and chips shop), or may include seating on-site (e.g. McDonalds and Burger King).
    - Narrow: any scope narrower than ‘moderate’. For example, definitions which include only chain/franchised fast food restaurants.
    - Broad: any scope broader than ‘moderate’, For example, definitions which include cafes, coffee shops, delis, bakeries/sandwich shops, sweet/desert stores, mobile takeaways.
- Convenience stores. The cut-points for convenience stores were selected to discriminate between broad classifications, which may include traditional convenience stores and also small grocery stores, moderate scope, which includes traditional convenience stores, and stores at gas stations, and narrow scope, which includes only traditional convenience stores.
    - Narrow: outlets falling under the scope of the NAICS classification for convenience stores (445120). Typically corresponds to outlets selling limited to no fresh produce.
    - Moderate: as for ‘narrow’, but also including gas stations.
    - Broad, includes traditional convenience stores and/or gas stations, and also small grocery stores, which may be required to have:
        - Annual sales < $1million
        - One cash register
        - Floor area < 15,000 square feet
        - Less than 3 employees
- Restaurants
    - Moderate: scope corresponding to NAICS code 722110 - restaurants primarily engaged in providing food services to patrons who order and are served while seated (i.e., waiter/waitress service) and pay after eating. Excludes bars/pubs, snack and non-alcoholic beverage (e.g. coffee) bars and theatres/entertainment venues.
    - Broad: includes other outlets e.g. coffee shops, snack bars, bars/pubs, entertainment venues.

Buffer Sizes

Buffer sizes used to compute access metrics were stratified as:

- ≤200m
- >200m - 400m (0.25 miles)
- >400m – 800m (0.5 miles)
- >800m – 1600m (1 mile)
- >1600m – 2400m (1.5 miles)
- 2400m – 3200m (2 miles)
- >3200m (2 miles)

These stratifications were selected to enable discrimination between commonly employed buffer sizes. For example, 400m/0.25mile, 800m/0.5mile, 1600m/1mile and 2 mile buffers are commonly used, and thus these distances were used as cut-points. Additionally, some studies used either 2km or 1.5 mile buffers, and thus an intermediate range of 1600m to 2400m (1.5 miles) was included. A small number of studies also used buffers ≤200m, and thus this category was added to capture those studies.

For alternative metrics, some measures were grouped as 'healthy', 'uhnealthy' and 'intermediate' metrics. Measures of the retail food environment were hypothesised to have a 'healthy', 'unhealthy', or 'intermediate' effect on obesity as set out below.

Healthy	Supermarkets Grocery stores
Unhealthy	Convenience stores Limited service/fast food restaurant Pizza establishments Bodegas
Intermediate	Small food store Full-service restaurant Total restaurants (including fast food) Prepared food site All food stores Independent/ethnic supermarket Food outlet diversity Living closer to healthy ethnic market than to healthy supermarket

Reporting of associations

Associations were only reported for methods or population groupings that were used in 5 or more studies to enable generalised comparisons between methods. Comparisons between methods or population groupings used in 4 or fewer studies were not made, because any differences observed were at increased likelihood of being due to differences in study context or specific design features, thereby confounding any observed patterns.

Supplement 7. List of included studies

Abbott G, Backholer K, Peeters A, Thornton L, Crawford D, Ball K. Explaining educational disparities in adiposity: The role of neighborhood environments. *Obesity (Silver Spring, Md)* 2014;**22:** 2413-2419.

An R, Sturm R. School and residential neighborhood food environment and diet among California youth. *Am J Prev Med* 2012;**42:** 129-135.

Bader MDM, Schwartz-Soicher O, Jack D, Weiss CC, Richards CA, Quinn JW*, et al.* More neighborhood retail associated with lower obesity among New York city public high school students. *Health and Place* 2013;**23:** 104-110.

Baek J, Hirsch JA, Moore K, Tabb LP, Barrientos-Gutierrez T, Lisabeth LD*, et al.* Statistical methods to study variation in associations between food store availability and body mass in the multi-ethnic study of atherosclerosis. *Epidemiology* 2017;**28:** 403-411.

Berge JM, Wall M, Larson N, Forsyth A, Bauer KW, Neumark-Sztainer D. Youth dietary intake and weight status: Healthful neighborhood food environments enhance the protective role of supportive family home environments. *Health Place* 2014;**26:** 69-77.

Block JP, Christakis NA, O'malley AJ, Subramanian S. Proximity to food establishments and body mass index in the Framingham Heart Study offspring cohort over 30 years. *Am J Epidemiol* 2011;**174:** 1108-1114.

Bloodworth RF, Ward KD, Relyea GE, Cashion AK. Food availability as a determinant of weight gain among renal transplant recipients. *Res Nurs Health* 2014;**37:** 253-259.

Bodicoat DH, Carter P, Comber A, Edwardson C, Gray LJ, Hill S*, et al.* Is the number of fast-food outlets in the neighbourhood related to screen-detected type 2 diabetes mellitus and associated risk factors? *Public Health Nutr* 2015;**18:** 1698-1705.

Bodor JN, Rice JC, Farley TA, Swalm CM, Rose D. The association between obesity and urban food environments. *Journal Of Urban Health: Bulletin Of The New York Academy Of Medicine* 2010;**87:** 771-781.

Boone-Heinonen J, Diez-Roux AV, Goff DC, Loria CM, Kiefe CI, Popkin BM*, et al.* The neighborhood energy balance equation: Does neighborhood food retail environment + physical activity environment = obesity? The CARDIA study. *PLoS One* 2013;**8:** e85141-e85141.

Brown AF, Vargas RB, Ang A, Pebley AR. The neighborhood food resource environment and the health of residents with chronic conditions. *J Gen Intern Med* 2008;**23:** 1137-1144.

Burdette HL, Whitaker RC. Neighborhood playgrounds, fast food restaurants, and crime: Relationships to overweight in low-income preschool children. *Prev Med* 2004;**38:** 57-63.

Burgoine T, Forouhi NG, Griffin SJ, Wareham NJ, Monsivais P. Associations between exposure to takeaway food outlets, takeaway food consumption, and body weight in Cambridgeshire, UK: Population based, cross sectional study. *BMJ* 2014;**348:** 1-10.

Carroll-Scott A, Gilstad-Hayden K, Rosenthal L, Peters SM, McCaslin C, Joyce R*, et al.* Disentangling neighborhood contextual associations with child body mass index, diet, and physical activity: The role of built, socioeconomic, and social environments. *Soc Sci Med* 2013;**95:** 106-114.

Casagrande SS, Franco M, Gittelsohn J, Zonderman AB, Evans MK, Fanelli Kuczmarski M*, et al.* Healthy food availability and the association with BMI in Baltimore, Maryland. *Public Health Nutr* 2011;**14:** 1001-1007.

Cerin E, Frank LD, Sallis JF, Saelens BE, Conway TL, Chapman JE*, et al.* From neighborhood design and food options to residents' weight status. *Appetite* 2011;**56:** 693-703.

Chen S, Florax RJ, Snyder S, Miller CC. Obesity and access to chain grocers. *Economic geography* 2010;**86:** 431-452.

Chen SE, Florax RJ, Snyder SD. Obesity and fast food in urban markets: A new approach using geo-referenced micro data. *Health Econ* 2013;**22:** 835-856 822p. Chaparro MP, Whaley SE, Crespi CM, Koleilat M, Nobari TZ, Seto E*, et al.* Influences of the neighbourhood food environment on adiposity of low-income preschool-aged children in Los Angeles county: A longitudinal study. *J Epidemiol Community Health* 2014;**68:** 1027-1033.

Chen D, Thomsen MR, Nayga RM, Jr., Bennett JL. Persistent disparities in obesity risk among public schoolchildren from childhood through adolescence. *Prev Med* 2016;**89:** 207-210.

Chen HJ, Wang Y. Changes in the neighborhood food store environment and children's body mass index at peripuberty in the United States. *J Adolesc Health* 2016;**58:** 111-118.

Cheng I, Shariff-Marco S, Koo J, Monroe KR, Yang J, John EM*, et al.* Contribution of the neighborhood environment and obesity to breast cancer survival: The California breast cancer survivorship consortium. *Cancer Epidemiol Biomarkers Prev* 2015;**24:** 1282-1290.

Dubowitz T, Ghosh-Dastidar M, Eibner C, Slaughter ME, Fernandes M, Whitsel EA*, et al.* The women's health initiative: The food environment, neighborhood socioeconomic status, BMI, and blood pressure. *Obesity (Silver Spring, Md)* 2012;**20:** 862-871.

Dubowitz T, Zenk SN, Ghosh-Dastidar B, Cohen DA, Beckman R, Hunter G*, et al.* Healthy food access for urban food desert residents: Examination of the food environment, food purchasing practices, diet and BMI. *Public Health Nutr* 2014;**18:** 2220-2230.

Dunn RA, Sharkey JR, Horel S. The effect of fast-food availability on fast-food consumption and obesity among rural residents: An analysis by race/ethnicity. *Econ Hum Biol* 2012;**10:** 1-13.

Fan JX, Hanson HA, Zick CD, Brown BB, Kowaleski-Jones L, Smith KR. 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. *BMJ Open* 2014;**4:** 1-9.

Fiechtner L, Block J, Duncan DT, Gillman MW, Gortmaker SL, Melly SJ*, et al.* Proximity to supermarkets associated with higher body mass index among overweight and obese preschool-age children. *Prev Med* 2013;**56:** 218-221.

Fiechtner L, Cheng ER, Lopez G, Sharifi M, Taveras EM. Multilevel correlates of healthy BMI maintenance and return to a healthy BMI among children in massachusetts. *Child Obes* 2017;**13:** 146-153.

Fiechtner L, Sharifi M, Sequist T, Block J, Duncan DT, Melly SJ*, et al.* Food environments and childhood weight status: Effects of neighborhood median income. *Child Obes* 2015;**11:** 260-268.

Ford PB, Dzewaltowski DA. Limited supermarket availability is not associated with obesity risk among participants in the Kansas WIC program. *Obesity* 2010;**18:** 1944-1951.

Ford PB, Dzewaltowski DA. Neighborhood deprivation, supermarket availability, and BMI in low-income women: A multilevel analysis. *Journal of Community Health: The Publication for Health Promotion and Disease Prevention* 2011;**36:** 785-796.

Galvez MP, Hong L, Choi E, Liao L, Godbold J, Brenner B. Childhood obesity and neighborhood food store availability in an inner city community. *Acad Pediatr* 2009;**9:** 339-343.

Gantner LA, Olson CM, Frongillo EA. Relationship of food availability and accessibility to women's body weights in rural upstate New York. *J Hunger Environ Nutr* 2013;**8:** 490-505.

Gary-Webb TL, Baptiste-Roberts K, Pham L, Wesche-Thobaben J, Patricio J, Pi-Sunyer FX*, et al.* Neighborhood and weight-related health behaviors in the look ahead (action for health in diabetes) study. *BMC Public Health* 2010;**10:** 11.

Gibson DM. The neighborhood food environment and adult weight status: Estimates from longitudinal data. *Am J Public Health* 2011;**101:** 71-78.

Gilliland JA, Rangel CY, Healy MA, Tucker P, Loebach JE, Hess PM*, et al.* Linking childhood obesity to the built environment: A multi-level analysis of home and school neighbourhood factors associated with body mass index. *Canadian Journal of Public Health* 2012;**103:** eS15-eS21.

Gose M, Plachta-Danielzik S, Willié B, Johannsen M, Landsberg B, Müller MJ. Longitudinal influences of neighbourhood built and social environment on children's weight status. *Int J Environ Res Public Health* 2013;**10:** 5083-5096.

Griffiths C, Frearson A, Taylor A, Radley D, Cooke C. A cross sectional study investigating the association between exposure to food outlets and childhood obesity in Leeds, UK. *Int J Behav Nutr Phys Act* 2014;**11:** 138.

Hattori A, An R, Sturm R. Neighborhood food outlets, diet, and obesity among California adults, 2007 and 2009. *Prev Chronic Dis* 2013;**10:** E35-E35.

Hickson DA, Diez Roux AV, Smith AE, Tucker KL, Gore LD, Lei Z*, et al.* Associations of fast food restaurant availability with dietary intake and weight among African Americans in the Jackson Heart Study, 2000-2004. *Am J Public Health* 2011;**101:** S301-S309.

Hobbs M, Green M, Griffiths C, Jordan H, Saunders J, McKenna J. How different data sources and definitions of neighbourhood influence the association between food outlet availability and body mass index: A cross-sectional study. *Perspect Public Health* 2017;**137:** 158-161.

Hosler AS, Michaels IH, Buckenmeyer EM. Food shopping venues, neighborhood food environment, and body mass index among guyanese, black, and white adults in an urban community in the US. *J Nutr Educ Behav* 2016;**48:** 361-368.e361.

Hutchinson PL, Nicholas Bodor J, Swalm CM, Rice JC, Rose D. Neighbourhood food environments and obesity in southeast Louisiana. *Health & Place* 2012;**18:** 854-860.

Inagami S, Cohen DA, Brown AF, Asch SM. Body mass index, neighborhood fast food and restaurant concentration, and car ownership. *J Urban Health* 2009;**86:** 683-695.

Jeffery RW, Baxter J, McGuire M, Linde J. Are fast food restaurants an environmental risk factor for obesity? *International Journal of Behavioral Nutrition and Physical Activity* 2006;**3:** 2.

Jiao J, Moudon AV, Kim SY, Hurvitz PM, Drewnowski A. Health implications of adults' eating at and living near fast food or quick service restaurants. *Nutr Diabetes* 2015;**5:** e171.

Jilcott Pitts SB, Keyserling TC, Johnston LF, Smith TW, McGuirt JT, Evenson KR*, et al.* Associations between neighborhood-level factors related to a healthful lifestyle and dietary intake, physical activity, and support for obesity prevention polices among rural adults. *J Community Health* 2015;**40:** 276-284.

Jilcott Pitts SB, Wu Q, McGuirt JT, Crawford TW, Keyserling TC, Ammerman AS. Associations between access to farmers' markets and supermarkets, shopping patterns, fruit and vegetable consumption and health indicators among women of reproductive age in eastern North Carolina, USA. *Public Health Nutr* 2013;**16:** 1944-1952 1949p.

Jilcott SB, Wade S, McGuirt JT, Qiang W, Lazorick S, Moore JB. The association between the food environment and weight status among eastern North Carolina youth. *Public Health Nutr* 2011;**14:** 1610-1617.

Jones-Smith JC, Karter AJ, Warton EM, Kelly M, Kersten E, Moffet HH*, et al.* Obesity and the food environment: Income and ethnicity differences among people with diabetes: The diabetes study of northern California (DISTANCE). *Diabetes Care* 2013;**36:** 2697-2705 2699p.

Kapinos KA, Yakusheva O, Eisenberg D. Obesogenic environmental influences on young adults: Evidence from college dormitory assignments. *Econ Hum Biol* 2014;**12:** 98-109.

Kestens Y, Lebel A, Chaix B, Clary C, Daniel M, Pampalon R*, et al.* Association between activity space exposure to food establishments and individual risk of overweight. *PLoS One* 2012;**7:** e41418.

Kruger DJ, Greenberg E, Murphy JB, DiFazio LA, Youra KR. Local concentration of fast-food outlets is associated with poor nutrition and obesity. *Am J Health Promot* 2014;**28:** 340-343.

Lamichhane AP, Puett R, Porter DE, Bottai M, Mayer-Davis EJ, Liese AD. Associations of built food environment with body mass index and waist circumference among youth with diabetes. *International Journal of Behavioral Nutrition and Physical Activity* 2012;**9:** 11.

Laraia BA, Downing JM, Zhang YT, Dow WH, Kelly M, Blanchard SD*, et al.* Food environment and weight change: Does residential mobility matter?: The diabetes study of northern California (DISTANCE). *Am J Epidemiol* 2017;**185:** 743-750.

Larsen K, Cook B, Stone MR, Faulkner GE. Food access and children's BMI in toronto, ontario: Assessing how the food environment relates to overweight and obesity. *Int J Public Health* 2015;**60:** 69-77.

Laska MN, Hearst MO, Forsyth A, Pasch KE, Lytle L. Neighbourhood food environments: Are they associated with adolescent dietary intake, food purchases and weight status? *Public Health Nutr* 2010;**13:** 1757-1763 1757p.

Laxy M, Malecki KC, Givens ML, Walsh MC, Nieto FJ. The association between neighborhood economic hardship, the retail food environment, fast food intake, and obesity: Findings from the survey of the health of Wisconsin. *BMC Public Health* 2015;**15:** 237.

Le H, Engler-Stringer R, Muhajarine N. Walkable home neighbourhood food environment and children's overweight and obesity: Proximity, density or price? *Can J Public Health* 2016;**107:** 5347.

Lee H. The role of local food availability in explaining obesity risk among young school-aged children. *Soc Sci Med* 2012;**74:** 1193-1203.

Leung CW, Laraia BA, Kelly M, Nickleach D, Adler NE, Kushi LH*, et al.* The influence of neighborhood food stores on change in young girls' body mass index. *Am J Prev Med* 2011;**41:** 43-51.

Li F, Harmer P, Cardinal BJ, Bosworth M, Johnson-Shelton D. Obesity and the built environment: Does the density of neighborhood fast-food outlets matter? *American journal of health promotion : AJHP* 2009a;**23:** 203-209.

Li F, Harmer P, Cardinal BJ, Bosworth M, Johnson-Shelton D, Moore JM*, et al.* Built environment and 1-year change in weight and waist circumference in middle-aged and older adults: Portland neighborhood environment and health study. *Am J Epidemiol* 2009b;**169:** 401-408.

Li F, Harmer PA, Cardinal BJ, Bosworth M, Acock A, Johnson-Shelton D*, et al.* Built environment, adiposity, and physical activity in adults aged 50-75. *Am J Prev Med* 2008;**35:** 38-46.

Li Y, Robinson LE, Carter WM, Gupta R. Childhood obesity and community food environments in Alabama's black belt region. *Child Care Health Dev* 2015;**41:** 668-676.

Liu GC, Wilson JS, Qi R, Ying J. Green neighborhoods, food retail and childhood overweight: Differences by population density. *Am J Health Promot* 2007;**21:** 317-325.

Lopez RP. Neighborhood risk factors for obesity. *Obesity (Silver Spring, Md)* 2007;**15:** 2111-2119.

Mejia N, Lightstone AS, Basurto-Davila R, Morales DM, Sturm R. Neighborhood food environment, diet, and obesity among Los Angeles county adults, 2011. *Prev Chronic Dis* 2015;**12:** E143.

Mellor JM, Dolan CB, Rapoport RB. Child body mass index, obesity, and proximity to fast food restaurants. *Int J Pediatr Obes* 2011;**6:** 60-68.

Miller LJ, Joyce S, Carter S, Yun G. Associations between childhood obesity and the availability of food outlets in the local environment: A retrospective cross-sectional study. *Am J Health Promot* 2014;**28:** e137-e145.

Minaker LM, Raine KD, Wild TC, Nykiforuk CIJ, Thompson ME, Frank LD. Objective food environments and health outcomes. *Am J Prev Med* 2013;**45:** 289-296.

Mobley LR, Root ED, Finkelstein EA, Khavjou O, Farris RP, Will JC. Environment, obesity, and cardiovascular disease risk in low-income women. *Am J Prev Med* 2006;**30:** 327-332.e321.

Moore K, Diez Roux AV, Auchincloss A, Evenson KR, Kaufman J, Mujahid M*, et al.* Home and work neighborhood environments in relation to body mass index: The multi-ethnic study of atherosclerosis (MESA). *J Epidemiol Community Health* 2013;**67:** 846-853.

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:** 333-339.

Morland KB, Evenson KR. Obesity prevalence and the local food environment. *Health & Place* 2009;**15:** 491-495.

Mushi-Brunt C, Haire-Joshu D, Elliott M, Brownson R. Fruit and vegetable intake and obesity in preadolescent children: The role of neighborhood poverty and grocery store access. *American Journal of Health Education* 2007;**38:** 258-265.

Nies MA, Weber KT, Holmes J, Peterson T, Serr K, Arias J*, et al.* Spatial and census data to evaluate obese persons and their environment (SCOPE). *Am J Health Behav* 2015;**39:** 582-588.

Ohri-Vachaspati P, DeLia D, DeWeese RS, Crespo NC, Todd M, Yedidia MJ. The relative contribution of layers of the social ecological model to childhood obesity. *Public Health Nutr* 2015;**18:** 2055-2066.

Ohri-Vachaspati P, Lloyd K, Delia D, Tulloch D, Yedidia MJ. A closer examination of the relationship between children's weight status and the food and physical activity environment. *Prev Med* 2013;**57:** 162-167 166p.

Oreskovic NM, Kuhlthau KA, Romm D, Perrin JM. Built environment and weight disparities among children in high- and low-income towns. *Acad Pediatr* 2009;**9:** 315-321.

Oreskovic NM, Winickoff JP, Kuhlthau KA, Romm D, Perrin JM. Obesity and the built environment among massachusetts children. *Clin Pediatr (Phila)* 2009;**48:** 904-912.

Pearson AL, Bentham G, Day P, Kingham S. Associations between neighbourhood environmental characteristics and obesity and related behaviours among adult New Zealanders. *BMC Public Health* 2014;**14:** 553.

Polsky JY, Moineddin R, Dunn JR, Glazier RH, Booth GL. Absolute and relative densities of fast-food versus other restaurants in relation to weight status: Does restaurant mix matter? *Prev Med* 2016;**82:** 28-34.

Pouliou T, Elliott SJ. Individual and socio-environmental determinants of overweight and obesity in urban Canada. *Health & Place* 2010;**16:** 389-398.

Powell LM, Han E. Adult obesity and the price and availability of food in the United States. *American Journal of Agricultural Economics* 2011;**93:** 378-384.

Prince SA, Kristjansso EA, Russell K, Billette J-M, Sawada MC, Ali A*, et al.* Relationships between neighborhoods, physical activity, and obesity: A multilevel analysis of a large Canadian city. *Obesity* 2012;**20:** 2093-2100.

Prince SA, Kristjansson EA, Russell K, Billette J-M, Sawada M, Ali A*, et al.* A multilevel analysis of neighbourhood built and social environments and adult self-reported physical activity and body mass index in Ottawa, Canada. *Int J Environ Res Public Health* 2011;**8:** 3953-3978.

Pruchno R, Wilson-Genderson M, Gupta AK. Neighborhood food environment and obesity in community-dwelling older adults: Individual and neighborhood effects. *Am J Public Health* 2014;**104:** 924-929.

Reitzel LR, Regan SD, Nga N, Cromley EK, Strong LL, Wetter DW*, et al.* Density and proximity of fast food restaurants and body mass index among African Americans. *Am J Public Health* 2014;**104:** 110-116 117p.

Richardson AS, Meyer KA, Howard AG, Boone-Heinonen J, Popkin BM, Evenson KR*, et al.* Multiple pathways from the neighborhood food environment to increased body mass index through dietary behaviors: A structural equation-based analysis in the CARDIA study. *Health Place* 2015;**36:** 74-87.

Rose D, Hutchinson PL, Bodor JN, Swalm CM, Farley TA, Cohen DA*, et al.* Neighborhood food environments and body mass index: The importance of in-store contents. *Am J Prev Med* 2009;**37:** 214-219.

Roth C, Foraker RE, Payne PRO, Embi PJ. Community-level determinants of obesity: Harnessing the power of electronic health records for retrospective data analysis. *BMC Med Inform Decis Mak* 2014;**14**.

Rummo PE, Guilkey DK, Ng SW, Meyer KA, Popkin BM, Reis JP*, et al.* Does unmeasured confounding influence associations between the retail food environment and body mass index over time? The coronary artery risk development in young adults (CARDIA) study. *Int J Epidemiol* 2017.

Rundle A, Neckerman KM, Freeman L, Lovasi GS, Purciel M, Quinn J*, et al.* Neighborhood food environment and walkability predict obesity in New York city. *Environ Health Perspect* 2009;**117:** 442-447.

Shariff-Marco S, Von Behren J, Reynolds P, Keegan TH, Hertz A, Kwan ML*, et al.* Impact of social and built environment factors on body size among breast cancer survivors: The pathways study. *Cancer Epidemiol Biomarkers Prev* 2017;**26:** 505-515.

Shier V, An R, Sturm R. Is there a robust relationship between neighbourhood food environment and childhood obesity in the USA? *Public Health* 2012;**126:** 723-730.

Shier V, Nicosia N, Datar A. Neighborhood and home food environment and children's diet and obesity: Evidence from military personnel's installation assignment. *Soc Sci Med* 2016;**158:** 122-131.

Spence JC, Cutumisu N, Edwards J, Raine KD, Smoyer-Tomic K. Relation between local food environments and obesity among adults. *BMC Public Health* 2009;**9:** 192-192.

Stark JH, Neckerman K, Lovasi GS, Konty K, Quinn J, Arno P*, et al.* Neighbourhood food environments and body mass index among New York city adults. *J Epidemiol Community Health* 2013;**67:** 736-742.

Sturm R, Datar A. Body mass index in elementary school children, metropolitan area food prices and food outlet density. *Public Health* 2005;**119:** 1059-1068.

Tamura K, Puett RC, Hart JE, Starnes HA, Laden F, Troped PJ. Spatial clustering of physical activity and obesity in relation to built environment factors among older women in three U.S. States. *BMC Public Health* 2014;**14:** 1322.

Thomsen MR, Nayga RM, Alviola PA, Rouse HL. The effect of food deserts on the body mass index of elementary schoolchildren. *American Journal of Agricultural Economics* 2016;**98:** 1-18.

Truong K, Fernandes M, An R, Shier V, Sturm R. Measuring the physical food environment and its relationship with obesity: Evidence from California. *Public Health* 2010;**124:** 115-118.

Tseng M, Thornton LE, Lamb KE, Ball K, Crawford D. Is neighbourhood obesogenicity associated with body mass index in women? Application of an obesogenicity index in socioeconomically disadvantaged neighbourhoods. *Health Place* 2014;**30:** 20-27.

Tung EL, Peek ME, Makelarski JA, Escamilla V, Lindau ST. Adult BMI and access to built environment resources in a high-poverty, urban geography. *Am J Prev Med* 2016;**51:** e119-e127.

Wall MM, Larson NI, Forsyth A, Van Riper DC, Graham DJ, Story MT*, et al.* Patterns of obesogenic neighborhood features and adolescent weight: A comparison of statistical approaches. *Am J Prev Med* 2012;**42:** e65-e75.

Wang MC, Kim S, Gonzalez AA, MacLeod KE, Winkleby MA. Socioeconomic and food-related physical characteristics of the neighbourhood environment are associated with body mass index. *J Epidemiol Community Health* 2007;**61:** 491-498.

Williams J, Scarborough P, Townsend N, Matthews A, Burgoine T, Mumtaz L*, et al.* Associations between food outlets around schools and BMI among primary students in England: A cross-classified multi-level analysis. *PLoS One* 2015;**10**.

Xu Y, Wen M, Wang F. Multilevel built environment features and individual odds of overweight and obesity in Utah. *Appl Geogr* 2015;**60:** 197-203.

Zenk SN, Mentz G, Schulz AJ, Johnson-Lawrence V, Gaines CR. Longitudinal associations between observed and perceived neighborhood food availability and body mass index in a multiethnic urban sample. *Health Educ Behav* 2017;**44:** 41-51.

Zhang YT, Laraia BA, Mujahid MS, Blanchard SD, Warton EM, Moffet HH*, et al.* Is a reduction in distance to nearest supermarket associated with BMI change among type 2 diabetes patients? *Health Place* 2016;**40:** 15-20.

Zhao Z, Kaestner R, Xu X. Spatial mobility and environmental effects on obesity. *Econ Hum Biol* 2014;**14:** 128-140.

Zick C, Smith KR, Fan JX, Brown BB, Yamada I, Kowaleski-Jones L. Running to the store? The relationship between neighborhood environments and the risk of obesity. *Soc Sci Med* 2009;**69:** 1493-1500.