

## Article

# Methods of measuring associations between the Retail Food Environment and weight status: Importance of classifications and metrics

Emma Wilkins<sup>a,\*</sup>, Michelle Morris<sup>b</sup>, Duncan Radley<sup>a</sup>, Claire Griffiths<sup>a</sup>

<sup>a</sup> Carnegie School of Sport, Leeds Beckett University, Leeds, UK

<sup>b</sup> Leeds Institute for Data Analytics, School of Medicine, University of Leeds, Leeds, UK

## ARTICLE INFO

## Keywords:

Community nutrition environment  
 Foodscape  
 GIS  
 Food access  
 Obesogenic environment  
 Methodology  
 United Kingdom

## ABSTRACT

Despite considerable research, evidence supporting associations between the 'Retail Food Environment' (RFE) and obesity remains mixed. Differences in the methods used to measure the RFE may explain this heterogeneity. Using data on a large ( $n = 10,111$ ) sample of adults from the Yorkshire Health Study (UK), we modelled cross-sectional associations between the RFE and weight status using (i) multiple definitions of 'Fast Food', 'Convenience' and 'Supermarkets' and (ii) multiple RFE metrics, identified in a prior systematic review to be common in the literature. Both the choice of outlet definition and the choice of RFE metric substantively impacted observed associations with weight status. Findings differed in relation to statistical significance, effect sizes, and directions of association. This study provides novel evidence that the diversity of RFE measurement methods is contributing to heterogeneous study findings and conflicting policy messages. Greater attention is needed when selecting and communicating RFE measures in research.

## 1. Introduction

The 'Retail Food Environment' (RFE) and its potential link with obesity has been the subject of considerable research for over a decade (Burdette & Whitaker, 2004; Cobb et al., 2015; Maddock, 2004). Interventions to create less 'obesogenic' RFEs have received notable policy interest across many western countries, with numerous local authorities now imposing restrictions on the RFE, such as the locations where 'unhealthy' food outlets can open (Diller & Graff, 2011; Local Government Association, 2016; Nykiforuk et al., 2018). The RFE comprises the availability and composition (e.g. relative mix) of food retailing within local environments and is thought to influence obesity by making 'unhealthy' foods more accessible and/or restricting access to 'healthy' foods. It may also influence obesity-related behaviours through more complex mechanisms, such as driving desire for foods, normalising behaviours, or establishing habits (Clary, Matthews, & Kestens, 2017). However, in spite of considerable research and numerous systematic reviews (Casey et al., 2014; Cobb et al., 2015; Feng, Glass, Curriero, Stewart, & Schwartz, 2010; Fleischhacker, Evenson, Rodriguez, & Ammerman, 2011; Gamba, Schuchter, Rutt, & Seto, 2015; Williams et al., 2014), international evidence supporting associations between the RFE and obesity is mixed. Null associations predominate in relation to the links between the RFE and obesity – making up around 75–80% of the literature (Wilkins et al., 2019). However, a notable

number of studies have found statistically significant associations of meaningful magnitude (Burgoine, Sarkar, Webster, & Monsivais, 2018; Dubowitz et al., 2012; Stark et al., 2013). For example, Burgoine, Forouhi (Burgoine et al., 2016) found that UK adults in the highest quartile of fast food exposure had a 0.9 units higher BMI than those in the lowest quartile. For an obese male of average height ( $BMI = 30 \text{ kg m}^{-2}$ , height = 178cm), a loss of 0.9 BMI units would equate to a weight loss of 2.9 kg, or ~3% body weight. This is a considerable difference, approaching the 5% considered to be clinically meaningful for weight-loss interventions by the National Institute for Health and Care Excellence; the UK body responsible for health care guidance (National Institute for Care Excellence, 2014).

One likely cause of the heterogeneity in outcomes is the diversity of methods used to measure the RFE. The majority (~two thirds) of international research (Caspi, Sorensen, Subramanian, & Kawachi, 2012; Lytle & Sokol, 2017), and numerous policy tools (Public Health England, 2016, 2017) use spatial measures to quantify the RFE. A previous review defined five 'dimensions of methodological diversity' across which spatial methods for measuring the RFE can differ: (i) choice of RFE data, (ii) data extraction methods, (iii) food outlet classifications, (iv) geocoding methods and (v) RFE measures (Wilkins et al., 2017a). Following from this, a systematic review of studies assessing associations between the spatially measured RFE and weight status in Western countries (Wilkins et al., 2019) (the 'Methods

\* Corresponding author. Leeds Beckett University, Headingley Campus, Headingley, Leeds, LS6 3QS, UK.

E-mail address: [e.wilkins@leedsbeckett.ac.uk](mailto:e.wilkins@leedsbeckett.ac.uk) (E. Wilkins).

Review') quantified, for the first time, RFE measurement methods across each of these dimensions, and showed that methods are highly varied. For example, while areal and buffer metrics (e.g. the count of outlets within buffers or administrative units) were used 242 times across the 113 studies, specific measures were used, at most, 15 times (count per area within 800 m - 1,600 m Euclidian buffers), and commonly no more than once.

Differences in measurement methods are likely to substantively impact study findings, and the corresponding policy messages. While most researchers acknowledge these differences, they still go on to make comparisons and draw conclusions across studies employing disparate methods. Little is known about the implications of using different spatial measurement methods, making interpretation and collation of the evidence challenging. Indeed, in spite of the high numbers of studies in this field, meta-analysis of the evidence has so far been precluded, with authors frequently citing the diversity of methods as a substantial hurdle to be overcome (Black et al., 2014a; Cobb et al., 2015; Williams et al., 2014). Ultimately, this is impeding translation of evidence into policy.

One aspect of methodological diversity that is not well understood is the choice of food outlet definitions. However, the Methods Review (Wilkins et al., 2019) found considerable variation in outlet definitions. For example, fast food outlets were sometimes defined narrowly as comprising only chain outlets, or in other cases defined broadly to include not only chain and non-chain 'traditional' fast food outlets but also establishments such as cafes and sandwich shops. Little is known about whether and to what extent differences in the definitions of commonly used outlet constructs impacts observed associations with weight status.

Another methodological aspect warranting investigation is the choice of RFE metric. This aspect was found to be the area of greatest methodological diversity in the Methods Review (Wilkins et al., 2019). Studies have used, for example, metrics of outlet count, count per population, count per area, presence/absence and various 'relative' metrics such as the ratio of fast food outlets to total outlets. There is some evidence from numerous countries including the US, Canada, Australia and the UK that differing RFE metrics lead to substantively different associations between the RFE and weight-related outcomes (Clary, Ramos, Shareck, & Kestens, 2015, 2016; Feng, Astell-Burt, Badland, Mavo, & Giles-Corti, 2018; Mason, Bentley, & Kavanagh, 2013; Polsky, Moineddin, Dunn, Glazier, & Booth, 2016; Shier, An, & Sturm, 2012; Thornton, Pearce, Macdonald, Lamb, & Ellaway, 2012). However, all existing studies investigate a limited selection of RFE metrics, which are not always relevant to existing literature. For example the two studies by Clary et al. (2015, 2016) compare relative and absolute kernel density metrics; types of metric that are rarely used in the literature. Thus, it remains unclear how the range of common RFE metrics compare. Further, many of these studies (Clary et al., 2015; Shier et al., 2012; Thornton et al., 2012) do not standardise RFE measures to allow direct comparison of effect sizes across measures with different units (e.g. count versus count per km<sup>2</sup>). Therefore, it is not possible to draw conclusions from these studies regarding the relative strength of associations.

This study sought to replicate common methods for measuring the RFE around the home, as identified in the Methods Review (Wilkins et al., 2019), and to compare associations between the RFE and weight status when employing:

- (i) different definitions of outlet constructs and
- (ii) differing RFE metrics.

Replication of common measurement methods ensured the findings of this present study were as relevant as possible to existing literature.

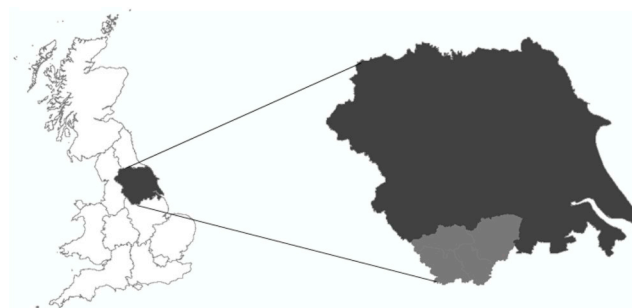


Fig. 1. Map of Britain showing location of the broad study area (Yorkshire & Humber – dark grey) and the initial recruitment area (South Yorkshire – light grey).

## 2. Methodology

### 2.1. Study sample

Data from the Yorkshire Health Study (YHS) were used to model associations between the RFE and weight status. The YHS is a longitudinal health survey of adult residents of the Yorkshire and Humber region of the UK. It originally focussed on recruiting residents of South Yorkshire, but was later expanded to cover the whole of the Yorkshire and Humber region (Fig. 1). Participants were invited by their medical practice to complete an online or paper-based survey on demographics, health conditions, behaviours, and health-care usage (response rate: 15.9%, further information at Supplement 1 and elsewhere (Green et al., 2014)). At the time of this study, two waves of measures were available: 2010–2012 and 2013–2015. Wave 1 comprised 27,806 participants. Of these, 10,876 (49.7%) also completed the wave 2 survey approximately 3 years later. The second wave sample additionally included 11,675 new participants.

To improve robustness of the data (in particular to spot inconsistencies in time-invariant variables between waves), the sample was restricted to only those with repeated measures ( $n = 10,876$ ), allowing validation of time-invariant measures such as gender (details below). However, all analyses were cross-sectional, using wave 1 data only. Longitudinal analyses were not performed, because the follow-up period was short, and the sample of insufficient size to detect differences in the trajectory of weight over time. This also replicates the majority (74%) of prior research, which also uses a cross-sectional design (Wilkins et al., 2019), allowing better generalisability to the extant literature.

Ethical approval for the original study was obtained from the Leeds East NHS Research Ethics (ref: 09/H1306/97). Ethical approval for the re-use of the YHS data in the present study was granted by the Local Research Ethics Committee at Leeds Beckett University.

### 2.2. Individual-level variables

From the YHS data, we used postcode of residence to calculate RFE exposures, and self-reported height and weight to derive outcomes of Body Mass Index (BMI, kg·m<sup>-2</sup>) and obesity classification (obese: BMI  $\geq 30$  kg m<sup>-2</sup>; non-obese: BMI  $< 30$  kg m<sup>-2</sup>). Self-reported age (years), gender (male/female), ethnicity (white/non-white) and highest education level ('less than high-school', 'high-school', 'college', or 'university', see Supplement 2) were used as covariates. Education level was used as a proxy for individual socioeconomic status (SES). These covariates were selected because they have been shown to be associated with weight status (Braveman, Cubbin, Egerter, Williams, & Pamuk, 2010; Health & Social Care Information Centre, 2017; Wang & Beydoun, 2007), and were found to be commonly controlled for the Methods Review (age: 90.3% of studies, gender: 89.4%, individual SES: 77.0%, ethnicity: 74.3%).

**Table 1**  
Food outlet classifications.

| Classification         | Definition  | Examples   |
|------------------------|---|--|
| Fast Food - Narrow     | Major chain outlets only.   | Comprised only of: McDonald's; Burger King; KFC; Domino's; Dixie Chicken; Wimpy; Chicken Cottage; Papa John's; Southern Fried Chicken (SFC); Five Guys; Harry Ramsdens; Subway; Little Chef. |
| Fast Food - Moderate   | 'Narrow' outlets, plus non-chain traditional fast food and takeaways with no/limited seating and no waitress service.   | Outlets serving burgers, kebabs, fried chicken, fish and chips, pizza, Indian and Chinese.   |
| Fast Food - Broad      | 'Moderate' outlets, plus takeaway cafes, retail bakeries and chain coffee shops.  | Subway, Starbucks, Costa Coffee, Greggs Bakery, Cooplands, Millie's Cookies, Shakeaway   |
| Convenience - Narrow   | Small convenience stores and newsagents, selling minimal or no fresh goods. Includes off-licenses and petrol station stores.  | Recognisable franchises include small-sized Mace, Londis, Costcutter, McColl's.  |
| Convenience - Moderate | 'Narrow' outlets, plus medium convenience stores selling a wider, but still limited range of fresh fruits and vegetables and frozen goods.                                | Recognisable franchises include medium-sized Nisa Local, Premier, Spar, McColl's, Londis, Costcutter.  |
| Convenience - Broad    | 'Moderate' outlets, plus grocery stores small enough to be exempt from the opening hour restrictions of the Sunday Trading Act 1994 (i.e. < 3000 square feet).            | Tesco Express/Metro, Sainsbury's Local, Co-operative (small stores), M&S Simply Food (small stores), large Nisa Local, Premier and Spar.   |
| Supermarket - Narrow   | Large chain supermarkets only. Often have long opening hours (e.g. 6am - 11pm or 24 h) and a wide range of facilities e.g. clothes/homeware departments.                  | Tesco, Morrisons, Sainsbury's, Asda or large Waitrose  |
| Supermarket - Moderate | 'Narrow' outlets, plus medium supermarkets with shorter opening hours and less extensive range of products and facilities. Large enough to fall under Sunday Trading Act. | Co-operative (large stores), M&S Simply Food (large stores), Waitrose (medium stores), Budgens, Aldi, Lidl, Iceland.   |
| Supermarket - Broad    | 'Moderate' outlets, plus small grocery stores, as defined for 'Convenience - broad' above.  | See 'Convenience - broad'.   |
| Restaurants            | Outlets serving evening meals and providing waited table service or a buffet.   | Pizza Express, Nandos, Zizzi's, La Tasca, Toby Carvery.  |
| Fruit and Veg Stores.  | Market stalls or outlets primarily retailing fruits and vegetables; including farm shops.   | 'Bob's Fruit and Veg', 'Hall's Green Grocers'  |
| Miscellaneous          | All other outlets - including cafes, pubs not serving food and speciality stores such as butchers and fishmongers.  | 'The Wrinkled Stocking Tea Room', 'Addys Butchers', 'Holland & Barrett', 'Cello Coffee House'.   |

### 2.3. Cleaning of YHS data

The range and distribution of the YHS data was checked for anomalies. We identified a number ( $n = 281$ ) of unexpected inconsistencies between measures of age, gender and height at baseline and follow-up. Thus, the sample was restricted to participants with repeated measures ( $n = 10,876$ ). This allowed use of the follow-up measures to exclude participants with unfeasible changes of gender ( $n = 30$ ), age ( $n = 61$ ), and height ( $n = 290$ ). Participants with missing baseline weight data ( $n = 277$ ), aged < 18 yrs at baseline ( $n = 66$ ) or located outside the study area ( $n = 1$ ) were also excluded. The final sample comprised 10,111 participants. Cleaning of the data did not notably change the sample demographics (Supplement 3).

### 2.4. Geocoding home locations

Participant home locations were geocoded to postcode centroids using ArcGIS v10.3.1 and a postcode lookup compiled from the Office for National Statistics postcode directory for the Yorkshire and Humber region (Ball, 2018). The postcodes were full unit postcodes, which contain an average of 15 addresses, with a maximum of 100 (Ordnance Survey, 2017). More precise home addresses were not available. The match rate was 100%.

### 2.5. Retail Food Environment measurement

This study sought to use the findings from the Methods Review (Wilkins et al., 2019) to replicate RFE measurement methods common in the literature. To supplement the description below, detailed rationale for the choice of methods, including data from the Methods Review on the prevalence of methods within the wider RFE-obesity literature can be found at Supplement 4.

#### 2.5.1. Retail Food Environment data

The Points of Interest (POI) dataset for 2011 (Ordnance Survey, 2011) was used to map food outlets. POI is produced by Ordnance Survey, the national mapping agency for Great Britain, and contains the

locations of a wide range of businesses and facilities. POI data were selected in preference to other RFE data for several reasons. First, it is a validated dataset, with very good spatial accuracy (all food outlets geocoded to points within building footprints), and comparable count accuracy to other UK and international RFE data (Wilkins et al., 2017b). Importantly, POI data were also available for the study period, which was not the case for other commonly used RFE data (e.g. food business data from local authorities). It is also frequently used in UK RFE research (Burgoine et al., 2009; Burgoine, Gallis, Penney, Monsivais, & Neelon, 2017; Cetateanu & Jones, 2014; Fraser, Clarke, Cade, & Edwards, 2012; Harrison et al., 2011; Jennings et al., 2011; Skidmore et al., 2010).

Food outlets were extracted from the POI dataset using proprietary classifications within the POI data. These included the POI classifications of 'restaurants', 'supermarkets', 'convenience stores and independent supermarkets', 'cafes, snack bars and tea rooms' and 'fast food and takeaway outlets' as well as more specialist outlets such as 'butchers' and 'bakeries' (complete list at Supplement 5).

Food outlets were mapped in ArcGIS v10.3.1 using the eastings and northings data within the POI dataset. Outlets falling within buffers relevant to this study (see below) were extracted, resulting in 5037 food outlets. Each outlet was then rigorously screened using a combination of automated and manual screening (details at Supplement 5) to remove 472 duplicates, 70 non-food outlets (e.g. wholesalers, private caterers, and distribution centres), and 26 outlets that had permanently closed prior to 2011. Several inaccuracies in easting and northing coordinates were identified and corrected ( $n = 21$ ), and a further eight outlets missing from the POI data were added. This resulted in a final sample of 4497 food outlets.

#### 2.5.2. Food outlet classification

Each food outlet was classified using a combination of (i) outlet name, (ii) proprietary classification, (iii) historic Google Street View images and (v) other information available online. This 'desk-based' method of classifying outlets has been previously shown to have good agreement (83%) with field-based food outlet classifications (Lake, Burgoine, Stamp, & Grieve, 2012). Outlets were assigned to at least one

of the following classifications: 'fast food', 'convenience', 'supermarket', 'restaurant', 'fruit & vegetable store' and 'miscellaneous' as defined in Table 1. These classifications were used to construct RFE exposures and covariates as described below.

The classifications 'fast food', 'convenience' and 'supermarket' were further divided into narrow, moderate and broad scopes (Table 1), reflecting definitions commonly employed in the RFE literature (Wilkins et al., 2019). Images of outlets falling within these classifications are shown at Supplement 6.

### 2.5.3. Aim 1: investigating outlet definitions

The first aim of this study was to investigate the impact of using different definitions of outlet constructs. We focussed on the most common constructs of 'fast food outlets', 'supermarkets' and 'convenience stores' (employed in 73.5%, 67.0% and 49.6% of studies respectively), which have notable heterogeneity in definitions. For this aim, the RFE was measured as counts of outlets within buffers, using network and Euclidian (circular) buffers around participant postcodes of residence, spanning 3,200 m, 1,600 m, 800 m and 400 m. These measures are among the most prevalent in RFE literature (Wilkins et al., 2019).

Network buffers used the oldest street network data available online from Ordnance Survey Open Roads (October 2016), including all roads except motorways. While this post-dates the POI and YHS data by 5 years, the street network is unlikely to have changed substantively during this period. Network buffers were generated using the Network Analyst Service Area tool in ArcGIS using generalised polygons with 100 m trim. Counts of narrow, moderate and broad definitions of 'Fast Food', 'Convenience' and 'Supermarkets' within the buffers were derived using the 'points in polygon' tool. Counts of other outlets (including 'Restaurants', 'Fruit & Veg' stores and 'Miscellaneous' outlets) within buffers were also calculated for use as covariates.

### 2.5.4. Aim 2: investigating RFE metrics

The second aim of this study was to investigate the impact of using different metrics to measure the RFE. To reduce the complexity of the results, this aspect focussed specifically on measures of 'Fast Food' within 800 m and 1,600 m buffers, which were found in the first part of the study (Aim 1), which is the most common construct employed in the literature (73.5% of studies). Network buffers were used because these allowed associations for raw counts and counts per area to be disentangled. 'Fast Food' outlets were defined using the moderate definition, which was found to be most prevalent in the literature (50.9% of studies (Wilkins et al., 2019)). Based on the Methods Review, we selected the following metrics for investigation:

1. Count of 'Fast Food' outlets;
2. Count of 'Fast Food' outlets per square kilometre;
3. Count of 'Fast Food' outlets per 1000 population;
4. Presence/absence of any 'Fast Food' outlets (binary variable);
5. The relative availability of outlet types, including:
  - a. 'Relative 1': Counts of ('Fast Food')/(Total Food Outlets);
  - b. 'Relative 2': Counts of ('Fast Food')/('Fast Food' + 'Restaurants');
  - c. 'Relative 3': Counts of ('Fast Food' + 'Convenience')/('Supermarkets' + 'Fruit & Veg Stores').

'Convenience' stores were defined using the 'broad' definition and 'Supermarkets' using the 'moderate' definition as defined in Table 1. 'Total Food Outlets' was the count of all food outlets, including 'Fruit & Veg' and 'Miscellaneous'.

### 2.6. Other environmental data

Several other sources of environmental data were used in this study (detailed information at Supplement 1). Firstly, the English Indices of Multiple Deprivation (IMD) for LSOAs (Department for Communities

and Local Government, 2015a) was used to control for area-level deprivation in our models. IMD is a government statistic which provides an indication of relative deprivation across England. The most recent IMD rankings (2015) were used because they were derived from measures predominantly collected in 2012/13 (Department for Communities and Local Government, 2015b), thus presenting the best possible temporal match to the YHS data. The 2011 Census Rural Urban Classifications for England at the LSOA level (Office for National Statistics, 2011a, 2011b) was also used to control for urbanicity. Each participant was assigned the IMD ranking and the urban/rural classification of the LSOA in which their residential postcode was located.

Finally, census data on the numbers of residents living within output areas across the study region in 2011 was obtained from the Office for National Statistics (Office for National Statistics, 2011a, 2011b) in order to calculate the number of food outlets per 1000 people. The populations within network buffers were estimated as the weighted sum of the populations living within output areas overlapping each buffer, with weightings corresponding to proportion of the output area overlapping the buffer. For example, if a buffer overlapped 40% of 'output area A' and 70% of 'output area B', the population of the buffer would be estimated as  $(0.4 \times P_A) + (0.7 \times P_B)$  where  $P_i$  is the population of output area  $i$ .

### 2.7. Statistical analyses

A series of regressions were run, which replicated models commonly used in the literature (Supplement 4). First, linear regressions were run to examine associations between each measure of the RFE and BMI (treated as a continuous variable). Corresponding logistic regressions were then run with the binary outcome of obesity ( $BMI \geq 30 \text{ kg m}^{-2}$  versus otherwise). In total 160 separate models were run:

- 132 models investigating the impact of construct definitions:
  - 'Fast Food' and 'Convenience' exposures: 16 models for each of the three outlet scopes and constructs respectively, the 16 models corresponding to all permutations of buffer size (3,200 m, 1,600 m, 800 m, 400 m), buffer type (Euclidian and network) and outcome (BMI and obesity).
  - 'Supermarket' exposures: 12 models for each of the three outlet scopes, the 12 models corresponding to all permutations of buffer size (3,200 m, 1,600 m, 800 m), buffer type (Euclidian and network) and outcome (BMI and obesity).
- 28 models investigating the impact of RFE metrics
  - 4 models for each of the 7 metrics, the 4 models corresponding to two buffer sizes (800 m network, 1600 m network) and two outcomes (BMI and obesity).

To allow for non-linear associations between RFE measures and BMI/obesity, RFE measures were modelled as quartiles where possible. This also allowed direct comparison of effect sizes for metrics with different units of measurement (e.g. counts versus counts per  $\text{km}^2$ ). Cut points were selected with the aim of having equal numbers of participants per quartile. However, due to the discretised nature of food outlet exposures, some quartiles were collapsed (Supplement 7). In particular, for the smallest buffer sizes, exposures were measured as presence/absence. Counts of 'Supermarkets' within 400 m buffers were not modelled due to insufficient variation (further details at Supplement 7).

Each measure of the RFE (i.e. each outlet definition, metric, and buffer size and type) was modelled separately, controlling for age (continuous), gender (binary), education (4 levels), ethnicity (white/non-white), neighbourhood deprivation (5 levels) and urbanicity (urban/rural). These variables were selected as they were found to be commonly controlled for in the Methods Review (Supplement 4). Models investigating the impact of outlet definitions additionally controlled for the count of all other outlets within buffers (modelled as quartiles). Just over half of all studies that provided clear descriptions

of statistical methods in our systematic review (Wilkins et al., 2019) included other outlets as covariates in models. Conversely, models investigating the impact of metric choice, *did not* control for other outlets, because inclusion of other outlets would lead to double-counting of the denominator for the relative metrics.

All statistical analyses were performed using R (version 3.2.3). The threshold for statistical significance was  $p < 0.05$ . Model fit was compared using Akaike Information Criterion (AIC) values and adherence to model assumptions was checked using diagnostic plots and generalised variance inflation factors standardised by degrees of freedom (GVIF<sup>1/(2DD)</sup>). We tested the sensitivity of our results to various alternative model parameterisations (treatment of RFE exposures as continuous and inclusion/exclusion of ‘other outlets’ as a covariate) and found no substantive differences in our main findings, although effect sizes and patterns of associations for specific RFE measures did differ.

### 3. Results

#### 3.1. Descriptive statistics

Descriptive statistics for the 10,111 YHS participants included in analyses are shown in Table 2. Compared to the English population in 2011, the YHS sample was older, with a higher proportion of white ethnicities and a small overrepresentation of females and people at the extremities of the education spectrum. Mean BMI was lower and there were fewer people classified as obese than among the general English population.

The majority ( $n = 9205$ ) of participants resided in South Yorkshire, which is an English metropolitan county spanning 1552 km<sup>2</sup>, with a population of 1.3 million (Office for National Statistics, 2016). A number of participants resided in the wider Yorkshire and Humber Region ( $n = 77$ ) and neighbouring parts of Derbyshire ( $n = 829$ ). Relative to England as a whole, participants tended to live in less deprived neighbourhoods (Table 2). There was nevertheless a good representation of the spectrum of deprivation within England, with participants living in LSOAs ranging from the 118th to the 32,309th most deprived (there being 32,844 LSOAs in England). Similar to the English population, the majority of participants resided in urban areas (population  $\geq 10,000$ ). The study area was, however, relatively unusual in that urban areas were predominantly designated as ‘minor conurbation’; collections of small towns and urban spaces that form a continuous urban area ( $n = 7750$ ; 91.0%). In contrast, across England only 3.6% of urban LSOAs have this designation.

Supplement 7 shows the distribution of RFE measures (e.g. food outlet counts) within buffers, together with correlations between measures. Most measures of the RFE exhibited a skewness to the right.

#### 3.2. Aim 1: investigating outlet definitions

Table 3 presents findings from all 132 models investigating the impact of construct definitions. The table summarises (i) whether models found statistically significant differences between quartile 1 and quartile 4, and (ii) the substantive conclusions supported by the models. It also shows the percentage agreement between the findings of models differing only in construct scope. Substantive conclusions took into account effect sizes and trends across quartiles in addition to p-values, and were classified as: ‘large positive/negative association’, ‘small positive/negative association’, ‘positive/negative U-shaped association’, and ‘null’ (see footnote of Table 3 for definitions). Further details regarding the classification of study findings can be found at Supplement 8.

Findings from models employing differing definitions for each outlet construct varied notably, both in terms of whether a statistically significant difference was observed, and in terms of the substantive conclusions drawn from the models. Agreement between the statistical significance of findings was particularly low for ‘Fast Food’

**Table 2**  
Sample characteristics.

| Variable                          | Level                       | Sample |                   | English Population <sup>a</sup> |
|-----------------------------------|-----------------------------|--------|-------------------|---------------------------------|
|                                   |                             | N      | %                 | %                               |
| <b>Individual Characteristics</b> |                             |        |                   |                                 |
| Gender                            | Female                      | 5749   | 56.9%             | 50.8%                           |
|                                   | Male                        | 4362   | 43.1%             | 49.2%                           |
| Age (years, range: 18–86)         | 18–29                       | 442    | 4.4%              | 20.7%                           |
|                                   | 30–39                       | 736    | 7.3%              | 16.9%                           |
|                                   | 40–49                       | 1418   | 14.0%             | 18.6%                           |
|                                   | 50–59                       | 2169   | 21.5%             | 15.3%                           |
|                                   | 60–69                       | 3126   | 30.9%             | 13.6%                           |
|                                   | 70–79                       | 1834   | 18.1%             | 8.9%                            |
|                                   | 80+                         | 386    | 3.8%              | 5.9%                            |
| Ethnic group                      | White                       | 9950   | 98.4%             | 85.4%                           |
|                                   | Non-white                   | 161    | 1.6%              | 14.6%                           |
| Education                         | < High school               | 4185   | 41.4%             | 35.8% <sup>b</sup>              |
|                                   | High school                 | 1383   | 13.7%             | 15.2% <sup>b</sup>              |
|                                   | A-level/similar             | 1673   | 16.5%             | 12.4% <sup>b</sup>              |
|                                   | University                  | 2871   | 28.4%             | 27.4% <sup>b</sup>              |
| BMI (kg·m <sup>-2</sup> )         |                             | 10,111 | 26.4 <sup>c</sup> | 27.1 <sup>c</sup>               |
| Weight Class                      | Obese                       | 1862   | 18.4%             | 25%                             |
|                                   | Non-obese                   | 8249   | 81.6%             | 75%                             |
| <b>Area Characteristics</b>       |                             |        |                   |                                 |
| Urbanicity                        | Urban                       | 8520   | 84.3%             | 82.4%                           |
|                                   | Rural                       | 1591   | 15.7%             | 17.6%                           |
| Area Deprivation                  | Quintile 1 (most deprived)  | 1818   | 18.0%             | 20.0%                           |
|                                   | Quintile 2                  | 1574   | 15.6%             | 20.3%                           |
|                                   | Quintile 3                  | 1596   | 15.9%             | 20.1%                           |
|                                   | Quintile 4                  | 2227   | 22.0%             | 19.9%                           |
|                                   | Quintile 5 (least deprived) | 2896   | 28.6%             | 19.7%                           |

N = number of participants.

<sup>a</sup> Obesity and BMI data from the Health Survey for England 2011. All other data from the 2011 Census for England.

<sup>b</sup> A further 9.3% of the UK population had a qualification classified as ‘other/apprenticeship’.

<sup>c</sup> Mean.

(31.3%–62.5% of models agreed), especially when contrasting narrow definitions to moderate and broad. Agreement between statistically significant findings from models employing different definitions of ‘Supermarkets’ and ‘Convenience’ was generally good (all  $> 80\%$ ). However, agreement decreased when considering substantive findings (50%–75%), suggesting that while definitions of ‘Supermarkets’ and ‘Convenience’ did not markedly affect statistical significance, they did impact inferences drawn from the data.

For all three constructs (‘Fast Food’, ‘Convenience’ and ‘Supermarkets’), AIC values did not differ markedly across the outlet scopes nor favour any particular scope (Supplement 8). The precision of point estimates (as indicated by 95% confidence intervals) also did not seem to be influenced by outlet scope, except in the smallest buffer sizes, where estimates for narrow definitions of ‘Fast Food’ and ‘Supermarkets’ were notably less precise.

Fig. 2 shows example associations between BMI and counts of ‘Fast Food’, ‘Convenience’ and ‘Supermarkets’ within 800–3200 m network buffers, according to the three definition scopes. All other exposures and outcomes are shown at Supplement 8.

#### 3.3. Aim 2: investigating RFE metrics

Table 4 summarises findings from the models investigating RFE metrics, with Fig. 3 illustrating associations between measures of ‘Fast Food’/‘Unhealthy Food’ exposure and BMI for 800 m and 1600 m buffers respectively. Results for models with obesity as an outcome are

**Table 3**  
Summary of findings for all 132 main models investigating the impact of outlet definition.

|                                 | Statistically significant difference (Q1 vs Q4) |    |                          |      |      | Substantive Conclusions |      |    |      |    |                          |      |      |      |  |
|---------------------------------|---|----|--------------------------|------|------|-------------------------|------|----|------|----|--------------------------|------|------|------|--|
|                                 | No. of associations                             |    | % agreement <sup>a</sup> |      |      | No. of associations     |      |    |      |    | % agreement <sup>a</sup> |      |      |      |  |
|                                 | +   | 0  | -                        | M    | N    | Lrg+                    | Sml+ | U+ | Null | U- | Sml-                     | Lrg- | M    | N    |  |
| <b>Fast Food<sup>b</sup></b>    |   |    |                          |      |      |                         |      |    |      |    |                          |      |      |      |  |
| B                               | 6   | 10 | 0                        | 62.5 | 43.8 | 3                       | 5    | 2  | 6    | 0  | 0                        | 0    | 68.8 | 25.0 |  |
| M                               | 8   | 8  | 0                        | -    | 31.3 | 2                       | 6    | 2  | 6    | 0  | 0                        | 0    | -    | 25.0 |  |
| N                               | 0   | 9  | 7                        | -    | -    | 0                       | 1    | 0  | 8    | 0  | 5                        | 2    | -    | -    |  |
| <b>Convenience<sup>b</sup></b>  |   |    |                          |      |      |                         |      |    |      |    |                          |      |      |      |  |
| B                               | 0   | 16 | 0                        | 87.5 | 87.5 | 0                       | 0    | 2  | 14   | 0  | 0                        | 0    | 56.3 | 62.5 |  |
| M                               | 1   | 14 | 1                        | -    | 87.5 | 0                       | 1    | 2  | 11   | 0  | 2                        | 0    | -    | 62.5 |  |
| N                               | 0   | 14 | 2                        | -    | -    | 0                       | 0    | 4  | 10   | 0  | 2                        | 0    | -    | -    |  |
| <b>Supermarkets<sup>c</sup></b> |   |    |                          |      |      |                         |      |    |      |    |                          |      |      |      |  |
| B                               | 1   | 11 | 0                        | 91.7 | 83.3 | 0                       | 1    | 3  | 8    | 0  | 0                        | 0    | 50.0 | 50.0 |  |
| M                               | 0   | 12 | 0                        | -    | 91.7 | 0                       | 0    | 2  | 10   | 0  | 0                        | 0    | -    | 75.0 |  |
| N                               | 0   | 11 | 1                        | -    | -    | 0                       | 0    | 0  | 10   | 0  | 2                        | 0    | -    | -    |  |

**Key:** + / - : statistically significant positive/negative difference between quartile 1 (Q1) and quartile 4 (Q4). 0: no statistically significant difference between Q1 and Q4. B: broad. M: moderate. N: narrow. Lrg + / - : large positive/negative association. Sml + / - : 'small positive/negative association' U + / - : 'positive/negative U-shaped association'.

<sup>a</sup> Percentage agreement between findings from models differing in the definition scope, but being otherwise identical.

<sup>b</sup> Results from 16 models respectively modelling exposures of outlet counts within 3,200m, 1,600m, 800m Euclidian and network buffers and presence/absence of outlets within 400m Euclidian and network buffers (8 exposures), against the respective outcomes of BMI and obesity (2 outcomes).

<sup>c</sup> Results from 12 models respectively modelling exposures of Supermarket counts within 3,200m & 1,600m Euclidian and network buffers and presence/absence of supermarkets within 800m Euclidian and network buffers (6 exposures), against the respective outcomes of BMI and obesity (2 outcomes).

shown at Supplement 9, and were substantively the same. Patterns of associations and effect sizes varied substantially. For example, the metrics Relative 1 ('Fast Food'/total outlets) and Relative 2 ('Fast Food'/'Fast Food' + 'Restaurants') consistently had the largest effect sizes, and exhibited dose-response relationships with weight status, with effect sizes trending upwards with increasing exposure. Conversely, for the count of 'Fast Food' outlets, all models revealed an inverted U-shaped association, wherein middle quartiles of exposure had statistically significantly higher BMI/obesity odds than the lowest quartile, but the highest and lowest quartiles were not different. The precision of point estimates (indicated by 95% confidence intervals) did not seem to be influenced by metric choice, except in 1,600 m buffers where the absolute count of 'Fast Food' was more precise than other metrics. Across the metrics, Relative 1 and Relative 2 generally had the lowest AIC values (Supplement 9) and thus the best model fit.

#### 4. Discussion

This study considers, for the first time, the impact of using different definitions for common outlet constructs when applied to a single dataset. It also expands upon emerging research into the choice of RFE metric by providing a comprehensive assessment of a range of metrics often used internationally in the literature. While it is unsurprising that we found measurement methods to influence findings to some degree, importantly we showed that the impact on findings was of material importance; influencing both statistical significance and substantive conclusions. Indeed, Fig. 4 demonstrates the extent to which findings differed when employing different measurement methods in this study. For each outlet type, models existed that supported associations in opposing directions. Methods used to measure the RFE are highly diverse (Wilkins et al., 2019) and this diversity undoubtedly contributes to the conflicting evidence base and confusing policy messages.

The definitions used to define outlet constructs were shown to impact findings; particularly for 'Fast Food', where associations for narrowly defined 'Fast Food' outlets were notably different from moderate

and broad definitions. This discrepancy is perhaps not surprising, given that the narrow definition omitted all non-chain hot food takeaways, which serve very similar food to chain fast food outlets, and were considerably more numerous than chain fast food outlets. Discrepancies in findings across definitions of 'Supermarkets' and 'Convenience' can also be explained in that these terms may encompass a range of outlets with different food offerings. Indeed, studies involving in-store audits have shown supermarkets and small food stores to have variable 'healthfulness' (encompassing measures such as healthy food availability, variety, price and promotions) (Black, Moon, & Baird, 2014b; Kersten, Laraia, Kelly, Adler, & Yen, 2012).

The importance of outlet definitions is often overlooked. The Methods Review upon which this study is based (Wilkins et al., 2019) identified that nearly half of all studies do not provide a clear definition for outlet constructs. Of those that did, over one quarter used a narrow definition similar to that used in the present study. Narrow definitions of fast food outlets produced markedly different results from moderate and broad scopes. Clear reporting of outlet definitions is therefore crucial to enable correct interpretation and translation of research into practice. Unsurprisingly, most existing systematic reviews do not differentiate between different definitions of common outlet constructs (Caspi et al., 2012; Cobb et al., 2015; Feng et al., 2010; Gamba et al., 2015). However, this may be hiding important associations.

There is currently little evidence to guide definition of food outlet constructs for the purposes of understanding RFE-obesity relationships. Moving forward, it is suggested that researchers use constructs that have policy-relevance, so that study findings are directly translatable to policy. For example, 'hot food takeaways' as defined within UK planning law is a distinct class of outlets against which policymakers can take action (e.g. by restricting where they can locate). Thus, within the UK context 'hot food takeaways' may present a preferred construct to measure.

Previous research has shown the specific metric used to measure the RFE can influence study findings (Clary et al., 2015, 2016; Feng et al., 2018; Mason et al., 2013; Polsky et al., 2016; Shier et al., 2012).

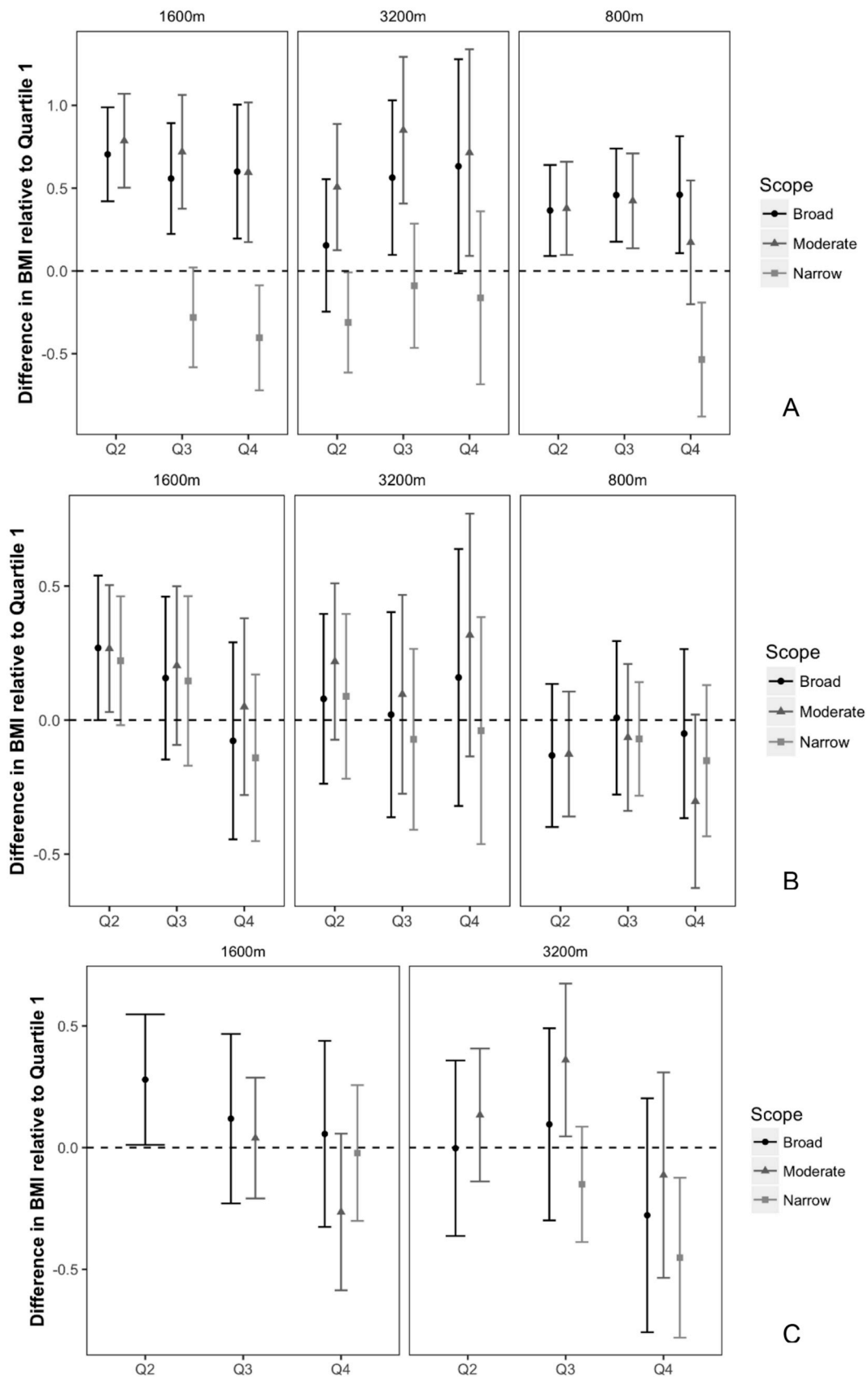


Fig. 2. Associations between counts of 'Fast Food' (A), 'Convenience' (B) and 'Supermarkets (C) and BMI within 800 m, 1600 m, and 3200 m network buffers. All models control for age, gender, ethnicity, education, urbanicity, area-level deprivation and count of all other food outlets within buffer. Values shown are beta coefficients representing the difference in BMI ( $\text{kg}\cdot\text{m}^{-2}$ ) between quartile 1 (reference category) and quartiles 2–4, and associated 95% confidence intervals. Point estimates are missing for some quartiles because these were collapsed with quartile 1 due to insufficient variation in outlet counts.

**Table 4**  
Summary of findings for all 28 main models investigating impact of metric choice.

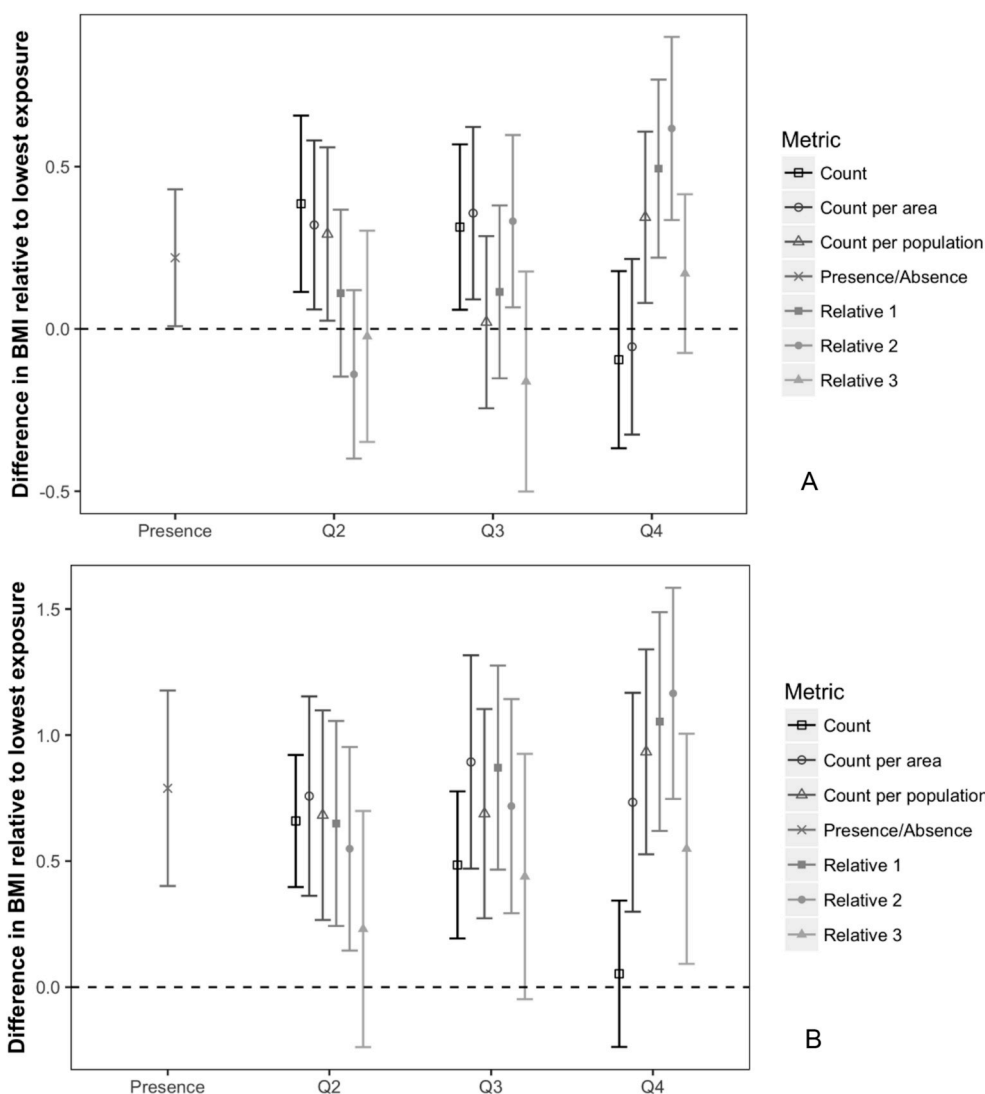
| Metric           | Statistically significant difference (Q1 vs Q4) <sup>a</sup> |   |   | Substantive conclusion <sup>a</sup> |       |     |      |    |      |      |
|------------------|--|---|---|-------------------------------------|-------|-----|------|----|------|------|
|                  | +  | 0 | - | Lrg +                               | Sml + | U + | Null | U- | Sml- | Lrg- |
| Count            | 0  | 4 | 0 | 0                                   | 0     | 4   | 0    | 0  | 0    | 0    |
| Count/Area       | 2  | 2 | 0 | 2                                   | 0     | 2   | 0    | 0  | 0    | 0    |
| Count/Population | 4  | 0 | 0 | 2                                   | 2     | 0   | 0    | 0  | 0    | 0    |
| Presence/Absence | 4  | 0 | 0 | 2                                   | 2     | 0   | 0    | 0  | 0    | 0    |
| Relative 1       | 4  | 0 | 0 | 4                                   | 0     | 0   | 0    | 0  | 0    | 0    |
| Relative 2       | 4  | 0 | 0 | 4                                   | 0     | 0   | 0    | 0  | 0    | 0    |
| Relative 3       | 2  | 2 | 0 | 0                                   | 2     | 0   | 2    | 0  | 0    | 0    |

**Key:** + / - : statistically significant positive/negative difference between quartile 1 (Q1) and quartile 4 (Q4). **0:** no statistically significant difference between Q1 and Q4. **B:** broad. **M:** moderate. **N:** narrow. **Lrg + /-:** large positive/negative association . **Sml + /-:** small positive/negative association **U + /-:** 'positive/negative U-shaped association .

<sup>a</sup> Results are from 4 models for each exposure, which respectively corresponded to 2 buffer sizes (1,600 m and 800 m) and two outcomes (BMI and obesity).

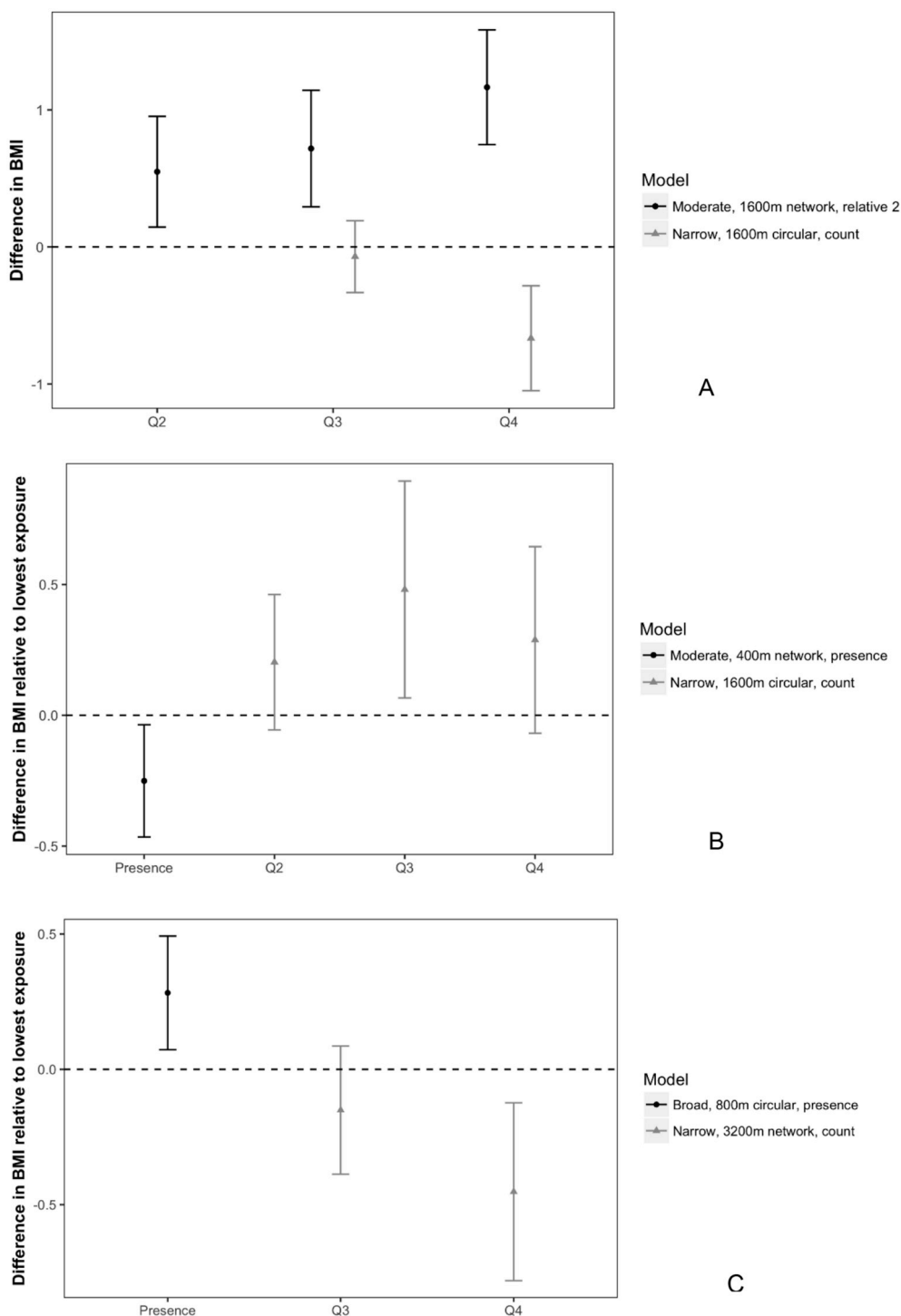
However, these studies have so far only investigated a limited range of metrics, which are not always relevant to existing literature. The present study is the first to investigate the impact of five of the most common types of area-based metric (count, count per area, count per

population, presence/absence and relative metrics), as identified a prior systematic review (Wilkins et al., 2019), allowing direct comparison across these key metrics. It showed that these metrics can lead to substantively different findings, even across conceptually similar



**Fig. 3.** Difference in BMI associated with increasing RFE exposure measures within 800 m (A) and 1600 m (B) network buffers. All models controlled for age, gender, ethnicity, education, urbanicity and area-level deprivation. Relative 1: fast food outlets/total outlets. Relative 2: fast food outlets/(fast food outlets + restaurants). Relative 3: (fast food outlets + convenience)/(supermarkets + fruit & veg stores). Values shown are beta coefficients representing the difference in BMI ( $\text{kg}\cdot\text{m}^{-2}$ ) between levels of exposure, relative to the lowest exposure (reference category), and associated 95% confidence intervals.





**Fig. 4. The most conflicting findings identified in this study in relation to the association between measures of Fast Food (A), Convenience (B), and Supermarket (C) exposures and BMI.** Values shown are beta coefficients representing the difference in BMI (kg-m<sup>-2</sup>) between quartile 1 (reference category) and quartiles 2–4 and associated 95% confidence intervals. Point estimates are missing for some quartiles because these were collapsed with quartile 1 due to insufficient variation in outlet counts.

metrics such as count and count per area. Prior systematic reviews have also tended to group multiple metrics together [e.g. 3, 9, 10], which may obscure important associations. Many papers also often interpret different metrics synonymously as indicating ‘access’ (Block, Christakis, O’Malley, & Subramanian, 2011; Bodor, Rice, Farley, Swalm, & Rose, 2010; Brown, Vargas, Ang, & Pebley, 2008; Chen, Thomsen, Nayga, & Bennett, 2016) or ‘availability’ (Baek et al., 2017; Bloodworth, Ward, Relyea, & Cashion, 2014; Casagrande et al., 2011). However, this may

give rise to over-simplistic messages to policymakers. Indeed, assuming RFE-obesity associations are causal (which is by no means definitive), our findings suggest that the specific aspect of the RFE that is regulated (e.g. the absolute number of outlets versus the relative availability of outlet types) would impact the effectiveness of interventions. Greater attention needs to be paid to RFE metrics moving forward.

The present study explored in detail three different types of relative metric. Relative metrics are increasingly being used in the RFE

literature because previous research has suggested they may lead to stronger and more consistent associations with weight-related outcomes than 'absolute' measures such as counts (Clary et al., 2015, 2016; Feng et al., 2018; Mason et al., 2013; Polsky et al., 2016; Shier et al., 2012). However, no previous research has compared multiple relative metrics to the range of absolute metrics commonly used in the literature. The present findings largely support previous studies that relative metrics give rise to larger effect sizes. The main models for 'Relative 1' (ratio of 'Fast Food' to total food outlets) and 'Relative 2' (ratio of 'Fast Food' to 'Fast Food' + restaurants) consistently had the largest effect sizes, supported clear, positive, dose-response associations with weight status, and had the best model fit. However, the present results also highlight the importance of the specific relative metric used; with associations for the 'Relative 3' metric (ratio of 'Fast Food' + 'Convenience' to 'Supermarkets' + 'Fruit & Veg') leading to smaller effect sizes and less consistent and clear associations with weight status. Thus, different relative metrics should not be interpreted synonymously.

The discrepancy in findings across metrics might be explained in that different metrics capture different dimensions of the RFE, acting through diverse causal pathways to drive obesity. It is often suggested that the RFE shapes obesity-related behaviours by facilitating or hindering direct access to foods. However, the RFE may also act through other pathways, such as stimulating desire for foods, normalising food behaviours, or building food habits (Clary et al., 2017). Measures of count per area or population may, for example, best capture the accessibility dimension. Conversely, the raw count of outlets (which measures absolute exposure) may better capture the desire dimension, and the relative availability of outlets (which might signal that eating a certain type of food is common or prevalent) may best capture the normalisation dimension. If this theory is correct, multiple measures would be necessary to adequately capture the RFE. The specific measures used would need to be considered in the context of the research question and the appropriate statistical frameworks and principles applied. While this study is focussed on spatial measures of the RFE it is also worth noting that there are also other non-spatial dimensions to the RFE, which may influence food purchasing behaviours, such as cost, opening hours, food quality and variety, and the degree to which outlets accommodate social or cultural requirements (Penchansky & Thomas, 1981).

An alternative explanation is that different metrics may be more strongly/weakly correlated with confounders of the RFE-obesity relationship, and thus models of RFE-obesity associations may be confounded to differing degrees, leading to differing patterns of association. In support of this theory, we found a notable number of inverted U-shaped relationships, particularly for measures of 'Fast Food' count. These U-shaped relationships may be explained by a latent variable that is positively correlated with 'Fast Food' count and negatively correlated with weight status. At high 'Fast Food' counts, the propensity of the latent variable to reduce weight status may dominate over the propensity of 'Fast Food' count to increase weight status, leading to a U-shaped relationship. A potential latent variable is street connectivity, which has been found to be associated with increased physical activity in several reviews and meta-analyses (Hajna et al., 2015; Saelens & Handy, 2008; Smith et al., 2017). Street connectivity is higher in urban areas and thus positively correlated with 'Fast Food' count, and may over-ride the influence of the RFE at high 'Fast Food' counts. Dividing the number of 'Fast Food' outlets by the number of total outlets would indirectly control for street connectivity, which might explain why the unexpected U-shaped relationship was replaced by a positive dose-response relationship for the ratio of 'Fast Food' to total outlets. Further research incorporating measures of potential environmental confounders is needed to corroborate this theory.

It is clear from the present study that different measures of the RFE should not be interpreted synonymously. However, given the diversity of measures used within the literature, this presents problems for the evaluation and collation of research. Moving forward, it is suggested

that researchers use multiple RFE measures (i.e. buffer sizes and metrics), ideally from a standardised set. A standardised set of metrics could be developed and agreed upon by a consortium of international experts, through consideration of existing measures and their theoretical underpinnings. It could include a range of standardised metrics (e.g. proximity, raw count, count relative to total outlets, presence/absence) and (where relevant) buffer sizes (e.g. 400 m, 800 m, 1600 m etc.) and types (e.g. network distances). Within these confines, researchers would then be able to select a range of metrics most suited to their research context. This would allow better comparability across studies, because overlap between studies in relation to the specific measures used would be more likely. It would also allow multiple aspects of the RFE to be captured and allow authors to draw more robust conclusions based on the findings from multiple measures rather than just a single measure. Standardised lists of measures already exist for certain specific contexts, such as that proposed by the Association of Public Health Epidemiologists in Ontario (Ahalya et al., 2017) for the reporting of public health statistics by public health units in Ontario. These could be used as a basis for the development of a more generalised set of standard measures.

This study has several strengths. First, results from a comprehensive systematic review were used to replicate and compare measures frequently used in RFE-obesity research. It is the first study to investigate the impact of using different definitions for outlet constructs, and also expands previous research into the impacts of metric choice, by considering the five most common types of area-based metric. The study used a large dataset from the YHS, and employed extensive data cleaning processes in respect of both YHS and POI data to ensure both datasets were as robust as possible. While this study uses a UK cohort, the methods investigated are used internationally, and therefore the main finding of this study – that different measurement methods used commonly across the international literature can lead to substantively different findings – has international relevance.

Limitations are that the YHS data were self-reported. Thus BMI is likely to be under-estimated; although self-reported data are still be useful for observing relationships in epidemiological research (Spencer, Appleby, Davey, & Key, 2007). The data cleaning process also identified a number of anomalies in the YHS data, such as changes in gender across waves. This casts doubt on the reliability of variables that could not be verified across waves. The YHS sample also differed from the English population across a number of characteristics. Of particular note, participants predominantly lived in urban areas designated as 'minor conurbations', which is a relatively unusual area designation in England. Relatedly, like the majority of RFE-obesity research, this study focussed solely on the home environment and neglected the numerous other environments that people are exposed to (Burgoine & Monsivais, 2013; Kwan, 2018). That said, given the focus of this study on the impact of methods, rather than inferences about the true associations between the RFE and weight status, these limitations are not of substantial importance. Indeed, the key finding of this study – that associations between the RFE and weight status depend on the measures used likely transcend population groups and environments.

Due to the discretised nature of food outlet counts, quartiles of exposure were not perfectly balanced and it was sometimes necessary to collapse quartiles to form a single category, which limits direct comparison of effect sizes. Stratification of exposures into quartiles also made comparison of associations more challenging, as each regression produced a parameter estimate for each quartile. Nevertheless, this parameterisation of the exposure was preferred over alternatives, such as the treatment of exposures as continuous linear variables, because it allowed for non-linear associations (which this study showed to be common), and also permitted better (albeit imperfect) comparability of effect sizes across metrics with differing units of measurement.

It was not the objective of this study to investigate causal relations between the RFE and weight status. This study used cross-sectional analyses, replicating common measurement and statistical methods

used in the literature, which may not necessarily be causally correct. Therefore, the relationships observed between the RFE and obesity cannot be interpreted as indicating causal associations. This study also does not provide evidence for the 'best' or 'correct' measures to use if seeking to understand causal associations. The observation that one measure gives rise to larger effect sizes does not mean that it is necessarily causally related to obesity. Furthermore, agreement between measures seemed to vary across buffer sizes and model specifications. Thus, the present findings are context specific and even in a predictive framework, we cannot conclude that any particular measure will be 'best' across all contexts.

## 5. Conclusion

This study highlights the important impact methodological decisions have on findings, and on the corresponding messages communicated to policymakers. Despite over a decade of research, evidence supporting associations between the RFE and obesity remains mixed, and this study provides evidence that methodological diversity is contributing to this mixed evidence base. Different measures of the RFE cannot be interpreted synonymously. However, the range of measures used in the literature is vast, presenting problems for the collation and interpretation of research. Moving forward, researchers should, at the very least, provide a clear and complete description of the measurement methods used. As there is little evidence to guide development of construct definitions, it is recommended that researchers either use definitions that are policy-relevant, as these will be more easily translatable into practice, or use definitions supported by well-articulated theory, so that clear hypotheses can be tested. It is also suggested that researchers use multiple RFE metrics, ideally from a standardised set of theory-informed metrics, to support comparability between studies, and allow more robust conclusions based on multiple, rather than a single metric.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Declarations of interest

None.

## Contributions

EW, MM, DR and CG were responsible for study conception and design. EW coordinated the review and acted as the primary reviewer. All other authors contributed as secondary reviewers. EW led the writing of the manuscript. All other authors provided critical feedback to shape the manuscript.

## Acknowledgements

We would like to acknowledge that The Yorkshire Health Study was funded by the NIHR CLAHRC Yorkshire and Humber ([www.clahrc-yh.nihr.ac.uk](http://www.clahrc-yh.nihr.ac.uk)). The views expressed are those of the author(s), and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. We would also like to thank Ordnance Survey for allowing us to use their data in this study.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ssmph.2019.100404>.

## References

- Ahalya, M., Jane, Y. P., Éric, R., Marc, L., Tina, M., & Leia, M. M. (2017). Geographic retail food environment measures for use in public health. *Health Promotion and Chronic Disease Prevention in Canada: Research, Policy and Practice*, 37(10), 357–362.
- Baek, J., Hirsch, J. A., Moore, K., Tabb, L. P., Barrientos-Gutierrez, T., Lisabeth, L. D., & Sánchez, B. N. (2017). Methods to study variation in the associations between food store availability and body mass in the multi-ethnic study of atherosclerosis. *Epidemiology*, 28(3), 403.
- Ball, C. (2018). *Yorkshire & the humber postcodes*. May 2018 Release . [cited 2018 16 June]; Available from: <https://www.doogal.co.uk/PostcodeDownloads.php>.
- Black, C., Moon, G., & Baird, J. (2014b). Measuring the healthfulness of food retail stores: Variations by store type and neighbourhood deprivation. *International Journal of Behavioral Nutrition and Physical Activity*, 11(69), 1–11.
- Black, C., Ntani, G., Inskip, H., Cooper, C., Cummins, S., Moon, G., & Baird, J. (2014a). Dietary inequalities: What is the evidence for the effect of the neighbourhood food environment? *Health & Place*, 27, 229–242.
- Block, J. P., Christakis, N. A., O'Malley, A. J., & Subramanian, S. V. (2011). Proximity to food establishments and body mass index in the Framingham heart study offspring cohort over 30 years. *American Journal of Epidemiology*, 174(10), 1108–1114.
- Bloodworth, R. F., Ward, K. D., Relyea, G. E., & Cashion, A. K. (2014). Food availability as a determinant of weight gain among renal transplant recipients. *Research in Nursing & Health*, 37(3), 253–259.
- Bodor, J. N., Rice, J. C., Farley, T. A., Swalm, C. M., & Rose, D. (2010). The association between obesity and urban food environments. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 87(5), 771–781.
- Braveman, P. A., Cubbin, C., Egerter, S., Williams, D. R., & Pamuk, E. (2010). Socioeconomic disparities in health in the United States: What the patterns tell us. *American Journal of Public Health*, 100(S1), S186–S196.
- Brown, A. F., Vargas, R. B., Ang, A., & Pebley, A. R. (2008). The neighborhood food resource environment and the health of residents with chronic conditions. *Journal of General Internal Medicine*, 23(8), 1137–1144.
- Burdette, H. L., & Whitaker, R. C. (2004). Neighborhood playgrounds, fast food restaurants, and crime: Relationships to overweight in low-income preschool children. *Preventive Medicine*, 38(1), 57–63.
- Burgoine, T., Lake, A. A., Stamp, E., Alvanides, S., Mathers, J. C., & Adamson, A. J. (2009). Research report: Changing foodscapes 1980–2000, using the ASH30 study. *Appetite*, 53(2), 157–165.
- Burgoine, T., Forouhi, N. G., Griffin, S. J., Brage, S., Wareham, N. J., & Monsivais, P. (2016). Does neighborhood fast-food outlet exposure amplify inequalities in diet and obesity? A cross-sectional study. *American Journal of Clinical Nutrition*, 103(6), 1540–1547.
- Burgoine, T., Gallis, J. A., Penney, T. L., Monsivais, P., & Neelon, S. E. B. (2017). Association between distance to nearest supermarket and provision of fruits and vegetables in English nurseries. *Health & Place*, 46, 229–233.
- Burgoine, T., Sarkar, C., Webster, C. J., & Monsivais, P. (2018). Examining the interaction of fast-food outlet exposure and income on diet and obesity: Evidence from 51,361 UK biobank participants. *International Journal of Behavioral Nutrition and Physical Activity*, 15(1), 71.
- Burgoine, T., & Monsivais, P. (2013). Characterising food environment exposure at home, at work, and along commuting journeys using data on adults in the UK. *International Journal of Behavioral Nutrition and Physical Activity*, 10(1), 85.
- Casagrande, S. S., Franco, M., Gittelsohn, J., Zonderman, A. B., Evans, M. K., Kuczmarski, M. F., & Gary-Webb, T. L. (2011). Healthy food availability and the association with BMI in Baltimore, Maryland. *Public Health Nutrition*, 14(6), 1001–1007.
- Casey, R., Oppert, J. M., Weber, C., Charreire, H., Salze, P., Badariotti, D., & Chaix, B. (2014). Determinants of childhood obesity: What can we learn from built environment studies? *Food Quality and Preference*, 31, 164–172.
- Caspi, C. E., Sorensen, G., Subramanian, S., & Kawachi, I. (2012). The local food environment and diet: A systematic review. *Health & Place*, 18(5), 1172–1187.
- Cetateanu, A., & Jones, A. (2014). Understanding the relationship between food environments, deprivation and childhood overweight and obesity: Evidence from a cross sectional England-wide study. *Health & Place*, 27, 68–76.
- Chen, D., Thomsen, M. R., Nayga, R. M., & Bennett, J. L. (2016). Persistent disparities in obesity risk among public schoolchildren from childhood through adolescence. *Preventive Medicine*, 89, 207–210.
- Clary, C. M., Ramos, Y., Shareck, M., & Kestens, Y. (2015). 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. *Preventive Medicine*, 71, 83–87.
- Clary, C., Lewis, D. J., Flint, E., Smith, N. R., Kestens, Y., & Cummins, S. (2016). The local food environment and fruit and vegetable intake: A geographically weighted regression approach in the ORiEL study. *American Journal of Epidemiology*, 184(11), 837–846.
- Clary, C., Matthews, S. A., & Kestens, Y. (2017). Between exposure, access and use: Reconsidering foodscape influences on dietary behaviours. *Health & Place*, 44, 1–7.
- Cobb, L. K., Appel, L. J., Franco, M., Jones-Smith, J. C., Nur, A., & Anderson, C. A. (2015). The relationship of the local food environment with obesity: A systematic review of methods, study quality, and results. *Obesity*, 23(7), 1331–1344.
- Department for Communities and Local Government (2015a). *English Index of multiple deprivation 2015*. UK Government. Retrieved from: <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015>, Accessed date: 1 February 2015.
- Department for Communities and Local Government (2015b). *The English Index of multiple deprivation (IMD) 2015* Technical Report . [cited 2018 1 October]; Available from:

- <https://www.gov.uk/government/publications/english-indices-of-deprivation-2015-technical-report>.
- Diller, P. A., & Graff, S. (2011). Regulating food retail for obesity prevention: How far can cities go? *Journal of Law Medicine & Ethics*, 39(s1), 89–93.
- Dubowitz, T., Ghosh-Dastidar, M., Eibner, C., Slaughter, M. E., Fernandes, M., Whitsel, E. A., & Escarce, J. J. (2012). The Women's Health Initiative: The food environment, neighborhood socioeconomic status, BMI, and blood pressure. *Obesity*, 20(4), 862–871.
- Feng, J., Glass, T. A., Curriero, F. C., Stewart, W. F., & Schwartz, B. S. (2010). The built environment and obesity: A systematic review of the epidemiologic evidence. *Health & Place*, 16(2), 175–190.
- Feng, X., Astell-Burt, T., Badland, H., Mavoa, S., & Giles-Corti, B. (2018). 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*, 49, 101–110.
- Fleischhacker, S., Evenson, K. R., Rodriguez, D. A., & Ammerman, A. S. (2011). A systematic review of fast food access studies. *Obesity Reviews*, 12(5), e460–e471.
- Fraser, L. K., Clarke, G. P., Cade, J. E., & Edwards, K. L. (2012). Fast food and obesity: A spatial analysis in a large United Kingdom population of children aged 13–15. *American Journal of Preventive Medicine*, 42(5), e77–e85.
- Gamba, R. J., Schuchter, J., Rutt, C., & Seto, E. Y. (2015). Measuring the food environment and its effects on obesity in the United States: A systematic review of methods and results. *Journal of Community Health*, 40(3), 464–475.
- Green, M. A., Li, J., Relton, C., Strong, M., Kearns, B., Wu, M., & Goyder, E. (2014). Cohort profile: The Yorkshire health study. *International Journal of Epidemiology*, 43(3), 707–712.
- Hajna, S., Ross, N. A., Brazeau, A.-S., Bélisle, P., Joseph, L., & Dasgupta, K. (2015). Associations between neighbourhood walkability and daily steps in adults: A systematic review and meta-analysis. *BMC Public Health*, 15(768), 1–8.
- Harrison, F., Jones, A. P., van Sluijs, E. M. F., Cassidy, A., Bentham, G., & Griffin, S. J. (2011). Environmental correlates of adiposity in 9–10 year old children: Considering home and school neighbourhoods and routes to school. *Social Science & Medicine*, 72(9), 1411–1419.
- Health & Social Care Information Centre (2017). Health survey for England 2016 adult overweight and obesity. [cited 2018 17 August]; Available from: <http://healthsurvey.hscic.gov.uk/media/63745/HSE2016-Adult-obe.pdf>.
- Jennings, A., Welch, A., Jones, A. P., Harrison, F., Bentham, G., van Sluijs, E. M. J., ... Cassidy, A. (2011). Local food outlets, weight status, and dietary intake: Associations in children aged 9–10 years. *American Journal of Preventive Medicine*, 40(4), 405–410.
- Kersten, E., Lارايا, B., Kelly, M., Adler, N., & Yen, I. H. (2012). Small food stores and availability of nutritious foods: A comparison of database and in-store measures, northern California, 2009. *Preventing Chronic Disease*, 9 E127–E127.
- Kwan, M.-P. (2018). The limits of the neighborhood effect: Contextual uncertainties in geographic, environmental health, and social science research. *Annals of the American Association of Geographers*, 108(6), 1482–1490.
- Lake, A. A., Burgoine, T., Stamp, E., & Grieve, R. (2012). The foodscape: Classification and field validation of secondary data sources across urban/rural and socio-economic classifications in England. *International Journal of Behavioral Nutrition and Physical Activity*, 9, 37–48.
- Local Government Association (2016). *Tipping the scales. Case studies on the use of planning powers to limit hot food takeaways*. London: Local Government Association.
- Lytle, L. A., & Sokol, R. L. (2017). Measures of the food environment: A systematic review of the field, 2007–2015. *Health & Place*, 44, 18–34.
- Maddock, J. (2004). The relationship between obesity and the prevalence of fast food restaurants: State-level analysis. *American Journal of Health Promotion*, 19(2), 137–143.
- Mason, K. E., Bentley, R. J., & Kavanagh, A. M. (2013). Fruit and vegetable purchasing and the relative density of healthy and unhealthy food stores – evidence from an Australian multilevel study. *Journal of Epidemiology & Community Health*, 67(3), 231–236.
- National Institute for Care Excellence (2014). *Weight management: Lifestyle services for overweight or obese adults*. [cited 2018 8 February]; Available from: <https://www.nice.org.uk/guidance/ph53/resources/weight-management-lifestyle-services-for-overweight-or-obese-adults-pdf-1996416726469>.
- Nykiforuk, C. I. J., Campbell, E. J., Macridis, S., McKennitt, D., Atkey, K., & Raine, K. D. (2018). Adoption and diffusion of zoning bylaws banning fast food drive-through services across Canadian municipalities. *BMC Public Health*, 18(1), 1–12.
- Office for National Statistics (2011). *English LSOA rural/urban classification 2011*. UK Data Service. Retrieved from: <http://census.ukdataservice.ac.uk/get-data/boundary-data.aspx>.
- Office for National Statistics (2012). *2011 Census: Population and household estimates for wards and output areas in England and Wales*. Office for National Statistics. Editor. 2012, Office for National Statistics. Retrieved from: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/>.
- Office for National Statistics (2016). *2011 census aggregate data (edition: February 2017)*. UK Data Service <https://doi.org/10.5257/census/aggregate-2011-2>.
- Ordnance Survey (2011). *Points of interest June 2011*. Ordnance Survey.
- Ordnance Survey (2017). *Postcode unit*. [cited 2017 30 October]; Available from: <http://data.ordnancesurvey.co.uk/ontology/postcode/PostcodeUnit>.
- Penchansky, R., & Thomas, J. W. (1981). The concept of access: Definition and relationship to consumer satisfaction. *Medical Care*, 19(2), 127–140.
- Polsky, J. Y., Moineddin, R., Dunn, J. R., Glazier, R. H., & Booth, G. L. (2016). Absolute and relative densities of fast-food versus other restaurants in relation to weight status: Does restaurant mix matter? *Preventive Medicine*, 82, 28–34.
- Public Health England (2016). *Obesity and the environment. Density of fast food outlets*. [cited 2017 13 April]; Available from: [https://www.noo.org.uk/securefiles/170413\\_1349/FastFoodmap\\_FINAL.pdf](https://www.noo.org.uk/securefiles/170413_1349/FastFoodmap_FINAL.pdf).
- Public Health England (2017). *Strategies for encouraging healthier 'out of home' food provision: A toolkit for local councils working with small food businesses*. [cited 2017 20 June]; Available from: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/604912/Encouraging\\_healthier\\_out\\_of\\_home\\_food\\_provision\\_toolkit\\_for\\_local\\_councils.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/604912/Encouraging_healthier_out_of_home_food_provision_toolkit_for_local_councils.pdf).
- Saelens, B. E., & Handy, S. L. (2008). Built environment correlates of walking: A review. *Medicine & Science in Sports & Exercise*, 40(7 Suppl), 550–566.
- Shier, V., An, R., & Sturm, R. (2012). Is there a robust relationship between neighbourhood food environment and childhood obesity in the USA? *Public Health*, 126(9), 723–730.
- Skidmore, P., Welch, A., Sluijs, E.v., Jones, A., Harvey, I., Harrison, F., ... Cassidy, A. (2010). Impact of neighbourhood food environment on food consumption in children aged 9–10 years in the UK SPEEDY (Sport, Physical Activity and Eating behaviour: Environmental Determinants in Young people) study. *Public Health Nutrition*, 13(7), 1022–1030.
- Smith, M., Hosking, J., Woodward, A., Witten, K., MacMillan, A., Field, A., ... Mackie, H. (2017). Systematic literature review of built environment effects on physical activity and active transport—an update and new findings on health equity. *International Journal of Behavioral Nutrition and Physical Activity*, 14(158), 1–27.
- Spencer, E. A., Appleby, P. N., Davey, G. K., & Key, T. J. (2007). Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutrition*, 5(4), 561–565.
- Stark, J. H., Neckerman, K., Lovasi, G. S., Konty, K., Quinn, J., Arno, P., ... Rundle, A. (2013). Neighbourhood food environments and body mass index among New York City adults. *Journal of Epidemiology & Community Health*, 67(9), 736–742.
- Thornton, L. E., Pearce, J. R., Macdonald, L., Lamb, K. E., & Ellaway, A. (2012). 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*, 11(29), 1–12.
- Wang, Y., & Beydoun, M. A. (2007). The obesity epidemic in the United States—gender, age, socioeconomic, racial/ethnic, and geographic characteristics: A systematic review and meta-regression analysis. *Epidemiologic Reviews*, 29(1), 6–28.
- Wilkins, E. L., Morris, M. A., Radley, D., & Griffiths, C. (2017a). Using geographic information systems to measure retail food environments: Discussion of methodological considerations and a proposed reporting checklist (Geo-FERN). *Health & Place*, 44, 110–117.
- Wilkins, E. L., Radley, D., Morris, M. A., & Griffiths, C. (2017b). Examining the validity and utility of two secondary sources of food environment data against street audits in England. *Nutrition Journal*, 16(82), 1–13.
- Williams, J., Scarborough, P., Matthews, A., Cowburn, G., Foster, C., Roberts, N., & Rayner, M. (2014). A systematic review of the influence of the retail food environment around schools on obesity-related outcomes. *Obesity Reviews*, 15(5), 359–374.
- Wilkins, E., Radley, D., Morris, M., Hobbs, M., Christensen, A., Lameck Marwa, A., Morrin, W., & Griffiths, C. (2019). A systematic review employing the GeoFERN framework to examine methods, reporting quality and associations between the retail food environment and obesity. *Health & Place*, 57, 186–199.