Title. The local food environment and fruit and vegetable intake: a Geographically Weighted Regression approach in The ORiEL Study

Short title. Spatial variations in environment-diet relationships

Authors. Christelle Clary, Daniel J. Lewis, Ellen Flint, Neil R. Smith, Yan Kestens, Steven Cummins

[1]Abbreviations

[2]Address for correspondence

Please cite as: Christelle Clary, Daniel J. Lewis, Ellen Flint, Neil R. Smith, Yan Kestens, Steven Cummins (in press) The local food environment and fruit and vegetable intake: a Geographically Weighted Regression approach in The ORiEL Study. *American Journal of Epidemiology*

Keywords: fruit and vegetable intake; geographically weighted regression; local food environment; non-stationarity; relative and absolute exposures

---

[1] Abbreviations: AICc, Akaike Information Criterion corrected; GWR, geographically weighted regression; ORiEL, Olympic Regeneration in East London.

[2] Christelle Clary, MSc, Department of Social and Preventive Medicine, ESPUM, Université de Montréal, 7101, avenue du Parc, Montréal (Québec) H3N 1X9, Canada (e-mail: christelle.clary@umontreal.ca; telephone: +0044 7455 054 868)

ABSTRACT

Studies that explore associations between the local food environment and diet routinely use global regression models, which assume that relationships are invariant across space. Yet, such stationarity assumptions have been little tested. We used global and geographically weighted regression models to explore associations between the residential food environment and fruit and vegetable intake. Analyses were performed in four East London boroughs, using the data of 969 adults in The ORiEL (Olympic Regeneration in East London) Study, collected between April and July 2012. Exposures were assessed both as absolute densities of healthy and unhealthy outlets taken separately, and as a relative measure (%healthy outlets). Overall, local models performed better than global models (decreased Akaike Information Criterion). Locally estimated coefficients varied across space, regardless of the type of exposure measure, although changes of sign were observed only when absolute measures were used. Despite findings from global models showing significant associations between the relative measure and fruit and vegetable intake (β=0.022, P<0.01) only, geographically weighted regression models using absolute measures outperformed those using relative measures. This study suggests that a greater attention should be given to non-stationary relationships between the food environment and diet. It further challenges the idea that one single measure of exposure, whether relative or absolute, can reflect the many ways the food environment may shape health behaviours.

Over the last decade, an extensive body of research has investigated how the local food environment may be related to dietary behaviours (1). A large majority of studies have used global regression to model the association between exposure to either healthy or unhealthy food environments and diet-related outcomes (2, 3). Reviews have however highlighted the lack of consistency in findings (2, 3), with associations being either positive (4-10), negative (11), or non-existent (12-18).

By using global regression models, researchers have implicitly relied on the assumption of a *stationary* relationship, that is, parameter estimates describe what is assumed to be an invariant relationship across space. However, public health researchers have begun to challenge the stationarity assumption. Using spatial regression modelling, such as geographically weighted regressions (GWR), a technique that allows for spatial variations in parameters estimates (19, 20), they have highlighted variations in associations across space between a range of environmental exposures and outcomes such as diet (21), obesity (22-27), active transportation (28), and birth weight (29). Fraser et al. have observed marked spatial variations, both in magnitude and nature, in the relationship between fast-food residential exposure and consumption among adolescents living in Bristol, United Kingdom (21). Higher exposure was significantly associated with increased consumption in some areas, but decreased consumption in others, even after adjusting for deprivation, gender, and physical activity levels. Overall, local modelling, when compared to global modeling, has shown better performance in terms of improved goodness-of-fit (22, 25, 27, 28), increased R^2 (22, 23, 25, 27, 28), and decreased spatial autocorrelation in regression residuals (22, 23, 25, 27)). Where non-stationarity has been pointed out, authors have therefore raised a caveat associated with the use of global models (22, 30).

Demonstrations of how a more spatially explicit approach to modelling might improve the understanding of associations between the food environment and diet is limited, however. To address this gap, this paper uses GWR models, alongside 'classical' global models, to predict fruit and vegetable intakes in relation to the residential food environment. Recent literature has suggested that relative measures, that account simultaneously for both healthy and unhealthy exposures, may be better correlates of diet than the traditional absolute measures, that account for either healthy or unhealthy outlets alone (31, 32). Therefore, residential exposure was assessed both as absolute densities of healthy and unhealthy outlets, taken separately, and as a relative measure as the percentage of healthy outlets.

METHODS

Data

Study design and participants. Adult socio-demographic and behavioural data were drawn from The ORiEL (Olympic Regeneration in East London) Study, a school-based longitudinal controlled quasi-experimental study in four boroughs of East London (Hackney, Tower Hamlets, Newham, and Barking and Dagenham) evaluating the health and social legacy of the London 2012 Olympics (33). For the purposes of the current study, baseline cross-sectional data collected between April 2012 and July 2012 on the parents of the adolescents in the surveyed schools were utilised (n=1277). Data were collected via computer assisted personal interviews with a response rate of 60%. Out of the 1277 interview participants, those whose home was located outside of the four study boroughs (n=31), who declared they had resided for less than a year at their current address (n=73), and for whom data was missing on the variables included in the models (n=228), were removed. 969 individuals were used for the analyses.

Food environment data. Data on the local food environment was collected from each of the four study boroughs, and in two coterminous boroughs in order to mitigate edge effects where the boundaries between study and non-study boroughs were most likely to be crossed by participants (Web Figure 1). It was assumed that the River Thames which bounds the southern study area acts as a natural barrier, and hence we did not collect data south of the river. Secondary data on the location of food businesses were obtained from the public register of food premises for each borough. 7,927 food outlets were geocoded to address-level using the OS MasterMap® Address Layer 2 product (Ordnance Survey, Southampton, UK) (34, 35). A sample was validated against concurrent Google Street View (Google Inc., Googleplex, Mountain View, California) and Bing Maps Streetside (Microsoft, Redmond, Washington) (see details about this validation in the Web Table 1). Finally, food businesses were classified according to a mutually exclusive classification informed by the literature (Web Table 1). Food outlet types were further categorized as "healthy" and "unhealthy". The term "healthy" referred to outlets that provide a range of fresh, frozen and canned fruits and vegetables options (6, 36), and included chain supermarkets, independent supermarkets, fruit and vegetable shops, and ethnic-specific supermarkets. Inversely, outlets providing few or no fruit and vegetable options were termed "unhealthy". They encompassed convenience stores, fast-food chains, and independent fast-food. The other listed outlet classes were not retained, as their affiliation to either “healthy” or “unhealthy” categories was rendered difficult by a lack of visibility regarding how many fruit and vegetable options they were assumed to offer.

Variables

Fruit and vegetable intake. Building on the Health Survey for England (http://data.gov.uk/dataset/health_survey_for_england), participants were asked two questions about their average daily portions - or handful size amount - of fruit and vegetable intake. Response categories were: “none”, “one”, “two”, “three”, “four”, and “five or more”. Fruits included fresh, frozen, canned and dried fruits, as well as fruit juices. Vegetables (excluding potatoes) encompassed fresh, frozen and canned vegetables. Fruit intake and vegetable intake were examined separately since previous research has found that the food environment may impact their intake differently (1, 5).

Food environment exposure. Kernel Density Estimations were computed in CrimeStat 3.3 (National Institute of JusticeUS) (37) for each of the seven food outlets categories retained. Kernel Density Estimation is an interpolation technique that transforms discrete spatial data into continuous density estimations based on a kernel of particular bandwidth and density function (38). The surface value is highest at the location of the observation point (i.e. food outlet location) and diminishes with increasing distance from this point, reaching its lowest value (e.g. zero for a quartic kernel) at the search radius (bandwidth) distance. The bandwidth can be fixed so that the distance from the observation point is either constant (*fixed* kernel density), or varies to maintain a constant number of observations under the curve (*adaptive* kernel density). The output is a raster file, with density estimates provided at each raster cell center, by adding the values of all kernel surfaces.

Density estimates were computed for raster cells of a 30 meter size, with a quartic function and an adaptive bandwidth using 7% of the nearest observations (Web Table 1). The use of adaptive rather than fixed bandwidths was motivated by two considerations. First, by accounting for the uneven distributions of food outlets, adaptive bandwidths avoid over- or under-smoothing of the continuous density function. The surface produced should thus be a better representation of the true density of the phenomenon than the more commonly used fixed bandwidth approach. Second, due to variations in the average distance between home and different outlet types, it is likely that people exhibit different spatial behaviours when accessing different types of food outlets (39-43). By allowing the bandwidth to vary as a function of the density of outlet types (larger and smaller bandwidths for more sparsely and more densely located outlet types, respectively), potential accessibility to these different outlet types may be better approximate. Several sizes of adaptive bandwidth were tested: 1%, 3%, 5%, 7%, and 10% of the nearest observations (Web Table 2). The 7% bandwidth was retained. Density values for each category of food outlets were extracted at participants' addresses, using the spatial analyst tools in ArcGIS 10.1 (Esri, Redlands, California).

Two absolute and one relative density-based measures of food environment exposure were computed at the participant's address level. Absolute exposure to healthy outlets was obtained by summing the densities of fruit and vegetable shops and chain, independent, and ethnic-specific supermarkets. Absolute exposure to unhealthy outlets was computed by summing the densities of convenience stores, fast-food chains, and independent fast-food restaurants. The proportion of healthy outlets was derived by dividing absolute exposure to healthy outlets by the sum of densities of health and unhealthy outlets.

Socio-demographic data. The following variables were considered as potential confounders of the relationship between the food environment and fruit and vegetable intake: age [continuous]; sex [female/male]; marital status [married, single, divorced/separated/widowed]; ethnic origin [White British/Irish, Asian/Asian British, Black/Black British, other]; highest qualification based on National Vocational Qualification level [none, low, intermediate, high, foreign]; and neighbourhood deprivation [low, medium-low, medium-high, high] based on the "income deprivation" score of the 2010 Index of Multiple Deprivation (Department for Communities and Local Government 2011) available for every Lower Layer Super Output Area (LSOA) in England. We extracted score values at participants' postcode using ArcGIS 10.1 and classified the resulting variable into quartiles. Because time spent at home may confound how the food environment relates to fruit and vegetable intake (44), we computed a proxy for time spent in the neighbourhood variable, termed time-budget [low, intermediate, high, other], derived from employment status and working hours. We considered time-budget to be inversely proportional to the time individuals reported spending in constrained activities, including work, studying and looking after the home or family members. Individuals working more than 30 hours a week were categorised as having a 'low' time budget; those working between 30 and 12 hours a week, being students, or looking after home or family were classified in the "intermediate" group. Finally, those working less than 12 hours a week, or being retired, or unemployed were considered having a 'high' time budget. An additional category - “other” - encompassed people with long-term sickness or disability, on maternity leave, on holiday or temporarily laid off, and individuals who declared another employment status.

Statistical analyses

First, fruit intake and vegetable intake were modelled separately using global linear regression (Ordinary Least Squares) in SPSS v.20 (IBM Corporation, Armonk, New York), with the three density measures used as the exposure variable in separate models. All six resulting models were fully adjusted for the following potential confounders: age, sex, marital status, qualification, ethnicity, time-budget, and neighbourhood deprivation. Because previous studies have found gender differences in the relationship between food environment exposure and dietary intake (15, 45-47), the interactions between sex and exposure were included in preliminary analyses, but excluded from each of the six models since they were not significant (p-values ranged from 0.141 to 0.972).

Second, spatial regressions (GWR) were performed with GWR 4.0.8 software (48), in order to account for the possible spatial non-stationarity of these relationships. Rather than calculating global parameter estimates based on one regression, GWR performs a series of local regressions with coefficients varying conditional on the location (i.e. participant address), drawing on the weighted surrounding data points (i.e. other participants’ location) (20). Only the β-coefficients of exposure to the food environment were allowed to locally vary over space, while the terms for other explanatory variables were fixed (semi-parametric GWR). We used a kernel with an adaptive Gaussian function, due to the uneven distribution of participants’ address (regression points), and with the bandwidth minimizing the corrected Akaike Information Criterion (AICc). The local β-coefficients of the six relationships under study, as well as the corresponding t-values, were mapped with ArcGIS 10.1 using Inverse Distance Weighted interpolations.

Third, AICc values, reported for both global regression and GWR models, were used to compare models' performance (49). The model with the lower AICc was taken as having a better fit. A difference in AICc of more than 3 values was regarded as a notable difference between two models (20). Eventually, spatial autocorrelation of standardized residuals was checked for both global and GWR models in GeoDa (GeoDa center, Arizona State University), using Moran'sI. Spatial weights were row-standardized and Euclidean inverse distance-based, with a 1 kilometer - bandwidth.

RESULTS

Descriptive analyses

Descriptive characteristics of participants are presented in Table 1. Participants were mostly women (75.5%), married (63.8%), and had predominantly no or low qualifications (22.7% and 29.2% respectively) and an intermediate time-budget (54.6%). The study sample was ethnically diverse with the largest group being "Asian/Asian British" (33.3%), followed by "White British/Irish" (25.0%), "Black/Black British" (22.5%) and "Other White Background" (13.8%). Three-quarters (74.6%) of participants declared eating two or more portions of fruits a day, 42.3% reported consuming three or more portions of vegetables, and 51.6% attained the recommended intake of five or more portions of fruits and vegetables (50). Participants' exposure to the food environment is presented in Table 2. On average, participants were exposed to 2.65 [Min: 0.13; Max: 15.03] healthy and 16.33 [1.26; 50.66] unhealthy outlets per km². The average percentage of healthy outlets around home was 12.55% [3.32%; 37.42%]. Participants with higher exposure to healthy outlets were more likely to also have higher exposure to unhealthy outlets (Table 2).

Global regressions

In fully-adjusted global models (Table 3), the proportion of healthy outlets was positively associated with both fruit intake (Model 3: β=0.022; $P < 0.01$) and vegetable intake (Model 6: β=0.022; $P < 0.01$). However, absolute measures of food environment exposure were not associated with any outcome. Both for fruit and vegetable intake, AICc values were lower for models using relative measures, suggesting better model performances. Spatial autocorrelation in standardised residuals was detected in all six models ($P<0.001$).

Local regressions

Table 3 shows that, compared to global regressions, local modelling was associated with both a decrease of AICc (around 12 points for models using relative exposures, up to 32 point for models using absolute exposure measures) and a suppression of spatial autocorrelation in standardised residuals. Furthermore, regression estimates varied across space (Figures 1 to 4). In the relationship between absolute measures and fruit or vegetable intake, there were spatial variations in the magnitude and sign of local estimates. Increased exposure to healthy outlets was significantly associated with increased fruit (Fig1A)&3A)) and vegetable (Fig2A)&4A)) intake in the central and extreme north-east parts of the study area, but with decreased vegetable intake in the south-western part. For unhealthy outlets, increased exposure was not significantly associated with decreased fruit intake (Fig1B)&3B)) and vegetable intake (Fig2B)&4B)) in the eastern part of the study area, but to increased fruit intake in the central part and increased vegetable intake in the eastern part of the study location. When relative measures were used, local estimates kept strictly positive, and were significant in the eastern half (vegetable intake – Fig2C)&4C)) and two-thirds (fruit intake – Fig1C)&3C)) of the study area. Local modelling of fruit intake performed the best (lowest AICc) when exposure was assessed as absolute density of unhealthy outlets. Local modelling of vegetable intake performed the best with exposure assessed as absolute density of healthy outlets.

DISCUSSION

This study used local (GWR) alongside global (Ordinary Least Squares) modelling to explore the relationship between different specifications of the food environment (absolute and relative measures of exposure) and dietary behaviours. We found evidence of non-stationarity in the relationship between the food environment and fruit and vegetable intake, regardless of the type of exposure measure. In line with previous studies (21-23, 25, 27), we observed that local regressions performed better than global regressions (decreased AICc and suppression of spatial autocorrelation in standardised residuals). Moreover, locally estimated coefficients and t-values varied across the study area. Among the plausible reasons for these variations, the omission of spatially structured determinants of fruit and vegetable intake is to be mentioned (20). Structural factors shaping individuals' behaviours across space (e.g. car ownership, proximity to public transport, walkable neighborhood) may modify how people relate to their residential environment (51, 52), by extending or conversely limiting food outlet experiences beyond the vicinity of home. In an attempt to account for the food environment individuals get exposed to away from home, all models were adjusted for a proxy of the time spent at home. This variable may, however, have only partially reflected the weight of non-residential exposures on the diet and residential food environment relationship. Additionally, personal income may not have been fully accounted for by using a proxy in the form of an aggregated measure of income at Lower Layer Super Output Area (LSOA) scale. Moreover, a few outlet types were excluded from exposure assessment (e.g. full-service restaurants), which may yet play a role in fruit and vegetable intake. As a result, potential misclassification issues are to be expected, especially in the western part of the study area where excluded outlets were particularly concentrated. Eventually, for models using absolute measures, considering only "healthy" outlets, while overlooking "unhealthy" options (and vice-versa) may plausibly give rise to an *omitted variable bias* (53). We found negative correlations ($P < 0.05$) between unhealthy (or healthy) outlet density and the locally estimated coefficients of the associations between healthy (or unhealthy, respectively) outlet density and fruit and vegetable intake (*results not shown*). Changes of sign across the study area when absolute measures are used may therefore partly result from not controlling for the alternate exposure. This finding suggests that greater consideration of relative measures of exposure is required, consistent with mounting evidence of association between relative measures of exposure and various health outcomes in the US (e.g. using the Retail Food Environment Index or RFEI (54-56)), Australia (32) and Canada (31, 57, 58).

The better performing (lower AICc) of the GWR models using absolute measures of exposure compared to those using the relative measure however warns us against the overly simplistic assertion that relative measures are “superior” to absolute measures. Exploration of potential threshold and saturation effects of relative measures of exposure on fruit and vegetable intake may help understand the better performing of local models using absolute measures. Masson et al. observed that, in Melbourne, for households in areas with between 10% and 15% of healthy outlets, the adjusted odds of healthier purchasing were 48% higher than those of households in areas with no more than 10% healthy food outlets (adjusted odds ratio (OR) =1.48, 95% confidence interval(): 1.12 to 1.96). However, the magnitude of the association for those living in areas with between 15% and 22% of healthy outlets was fairly similar to that estimated for the middle category of the relative measure (i.e. 10–15%). Mason et al.’s findings suggest a possible saturation effect above 10% of healthy stores. Furthermore, estimates failed to reach significance in areas located in the western part of the study, regardless of the type of exposure measure. An explanation may be that, in environments abundantly provided with food sources (which is the case in the western part of the study), the influence of exposure to the food environment may decline in favour of socio-economic or personal influences (59).

Alongside model misspecifications, non-stationarity may therefore also stem from *intrinsic* differences in the way individuals respond to specific characteristics (i.e. densities of healthy outlets or unhealthy outlets, or relative densities of healthy and unhealthy outlets) of their local food environment (20). The relationship of individuals to their environment is likely to be of a recursive nature, with environments and individuals mutually influencing each other (60, 61). Over time, this recursive relationship may induce specific emerging properties pushing individuals to relate differently to some specifically measured characteristics of their environment. Identifying the contextual conditions under which absolute and relative measures of exposure should be used to model fruit and vegetable intake would help better quantify the influence of local food environments on dietary behaviours. In turn, this would help develop locally-targeted policy interventions accounting for context-specificities.

Limitations

First, the cross-sectional nature of our data precludes any conclusions regarding causal relationships between exposure to the food environment and fruit and vegetable intakes. Second, external validity of this study is limited since ORiEL participants are non-representative of the United Kingdom population as whole. The overrepresentation of women (75.5%), and some ethnic minorities in our sample may have impacted the declared intake of fruit and vegetable of participants (5, 62, 63), which is approximately twice as great as the national figure (62). However, our findings are conceptually acceptable and consistent with existing literature (21, 31, 32). Third, due to ambiguities over how to classify a few food outlet types as "healthy" and "unhealthy", we excluded from our analysis outlets that may have yet played an important role in fruit and vegetable intake. Not fully accounting for the local food environment may have resulted in biased estimates. Fourth, exposure measures derived by using an adaptive bandwidth do not account for possible individualised spatial mobility potential (e.g. having access to a car) (64). Better exposure assessment would benefit from investigating individuals' movements across space and time, and considering individuals' perception of the food environment (65) – for example through using ecological momentary assessments (EMAs) (66). Finally, multi-collinearity issues among local coefficients have been raised in GWR models (67). However, this is unlikely to have a major impact on our findings as only the β-coefficients of exposure to the food environment were allowed to vary over space.

Conclusion

In this paper we demonstrated that global models using the relative measure outperformed those using absolute measures. The local modelling of exposure-diet relationships, however, suggests that the relative measure may not fully capture the complexity of environmental risks for dietary behaviours. We have highlighted spatial variations in the association between the food environment and fruit and vegetable intake, regardless of the type of exposure measure. Moreover, local models using absolute measures outperformed those using the relative measure. More research on the relative contributions of absolute and relative measures of exposure is needed in order to guide efficient contextual policy responses to unhealthy dietary behaviours. GWR can be a useful tool to better understand the role of contextual factors that shape how individuals respond to their local food environment.

Acknowledgements

Author affiliations. Department of Social and Preventive Medicine, Université de Montréal, Montreal, Canada (Christelle Clary, Yan Kestens); CRCHUM – Montreal University Hospital Research Centre, Montréal, Canada (Christelle Clary, Yan Kestens); Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, London, UK (Daniel J. Lewis, Ellen Flint, Steven Cummins); Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK (Neil R. Smith).

Funding sources. The ORiEL Study (www.orielproject.co.uk) is funded by the National Institute for Health Research (NIHR) Public Health Research Programme (Grant number: 09/3005/09). Steven Cummins is supported by a United Kingdom National Institute of Health Research Senior Fellowship. Yan Kestens holds a CIHR Chair in Urban Interventions and Population Health. Christelle Clary is supported by the Research Centre, Centre Hospitalier de l'Université de Montréal (CRCHUM).

Thank-yous. This paper was written while Christelle Clary was a visiting student at the London School of Hygiene and Tropical Medicine (LSHTM). She would like to thank the LSHTM, and more specifically the Professor Steven Cummins and his team, for hosting her visit, making The ORiEL data available, and feeding the thoughts that allowed for the production of the present paper.

Conflict of interest. The authors have no conflict of interest to declare.

REFERENCES

1. Kamphuis CB, Giskes K, de Bruijn GJ, et al. Environmental determinants of fruit and vegetable consumption among adults: a systematic review. *Brit J Nutr* 2006;96(4):620-635.
2. Charreire H, Casey R, Salze P, et al. Measuring the food environment using geographical information systems: a methodological review. *Public Health Nutr* 2010;13(11):1773-1785.
3. Caspi CE, Sorensen G, Subramanian SV, et al. The local food environment and diet: a systematic review. *Health Place* 2012;18(5):1172-1187.
4. Zenk SN, Lachance LL, Schulz AJ, et al. Neighborhood retail food environment and fruit and vegetable intake in a multiethnic urban population. *Am J Health Promot* 2009;23(4):255-264.
5. Bodor JN, Rose D, Farley TA, et al. Neighbourhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. *Public Health Nutr* 2008;11(4):413-420.
6. Rose D, Richards R. Food store access and household fruit and vegetable use among participants in the US Food Stamp Program. *Public Health Nutr* 2004;7(8):1081-1088.
7. Zenk SN, Schulz AJ, Hollis-Neely T, et al. Fruit and vegetable intake in African Americans income and store characteristics. *Am J Prev Med* 2005;29(1):1-9.
8. Laraia BA, Siega-Riz AM, Kaufman JS, et al. Proximity of supermarkets is positively associated with diet quality index for pregnancy. *Prev Med* 2004;39(5):869-875.
9. Moore LV, Diez Roux AV, Nettleton JA, et al. Fast-food consumption, diet quality, and neighborhood exposure to fast food: the multi-ethnic study of atherosclerosis. *Am J Epidemiol* 2009;170(1):29-36.
10. Satia JA, Galanko JA, Siega-Riz AM. Eating at fast-food restaurants is associated with dietary intake, demographic, psychosocial and behavioural factors among African Americans in North Carolina. *Public Health Nutr* 2004;7(8):1089-1096.
11. Asfaw A. Does supermarket purchase affect the dietary practices of households? Some empirical evidence from Guatemala. *Development Policy Review* 2008;26(2):227-243.
12. Cummins S, Petticrew M, Higgins C, et al. Large scale food retailing as an intervention for diet and health: quasi-experimental evaluation of a natural experiment. *J Epidemiol Community Health* 2005;59(12):1035-1040.
13. Pearce J, Hiscock R, Blakely T, et al. The contextual effects of neighbourhood access to supermarkets and convenience stores on individual fruit and vegetable consumption. *J Epidemiol Community Health* 2008;62(3):198-201.
14. Jeffery RW, Baxter J, McGuire M, et al. Are fast food restaurants an environmental risk factor for obesity? *Int J Behav Nutr Phys Act* 2006;3(2):1-6.
15. Macdonald L, Ellaway A, Ball K, et al. Is proximity to a food retail store associated with diet and BMI in Glasgow, Scotland? *BMC public health* 2011;11(464):1-9.
16. Zenk SN, Schulz AJ, Matthews SA, et al. Activity space environment and dietary and physical activity behaviors: a pilot study. *Health Place* 2011;17(5):1150-1161.
17. Pearce J, Hiscock R, Blakely T, et al. A national study of the association between neighbourhood access to fast-food outlets and the diet and weight of local residents. *Health Place* 2009;15(1):193-197.
18. Cummins S, Flint E, Matthews SA. New neighborhood grocery store increased awareness of food access but did not alter dietary habits or obesity. *Health Affair* 2014;33(2):283-291.

19. Fotheringham AS, Charlton ME, Brunsdon C. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environ Plann A* 1998;30(11):1905-1927.
20. Fotheringham AS, Brunsdon C, Charlton ME eds. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Chichester, UK: Wiley, 2002.
21. Fraser LK, Clarke GP, Cade JE, et al. Fast food and obesity: a spatial analysis in a large United Kingdom population of children aged 13-15. *Am J Prev Med* 2012;42(5):77-85.
22. Chi S-H, Grigsby-Toussaint DS, Bradford N, et al. Can geographically weighted regression improve our contextual understanding of obesity in the US? Findings from the USDA Food Atlas. *Appl Geogr* 2013;44:134-142.
23. Chalkias C, Papadopoulos AG, Kalogeropoulos K, et al. Geographical heterogeneity of the relationship between childhood obesity and socio-environmental status: Empirical evidence from Athens, Greece. *Appl Geogr* 2013;37:34-43.
24. Procter KL, Clarke GP, Ransley JK, et al. Micro-level analysis of childhood obesity, diet, physical activity, residential socioeconomic and social capital variables: where are the obesogenic environments in Leeds? *Area* 2008;40(3):323-340.
25. Chen D-R, Truong K. Using multilevel modeling and geographically weighted regression to identify spatial variations in the relationship between place-level disadvantages and obesity in Taiwan. *Appl Geogr* 2012;32(2):737-745.
26. Black NC. An ecological approach to understanding adult obesity prevalence in the United States: A county-level analysis using geographically weighted regression. *Applied Spatial Analysis and Policy* 2014;7(3):283-299.
27. Xu Y, Wang F. Built environment and obesity by urbanicity in the U.S. *Health Place* 2015;34:19-29.
28. Feuillet T, Charreire H, Menai M, et al. Spatial heterogeneity of the relationships between environmental characteristics and active commuting: towards a locally varying social ecological model. *Int J Health Geogr* 2015;14(12):1-14.
29. Tu J, Tu W, Tedders SH. Spatial variations in the associations of birth weight with socioeconomic, environmental, and behavioral factors in Georgia, USA. *Appl Geogr* 2012;34:331-344.
30. Nakaya T, Fotheringham AS, Brunsdon C, et al. Geographically weighted Poisson regression for disease association mapping. *Stat Med* 2005;24:2695-2717.
31. Clary CM, Ramos Y, Shareck M, et al. Should we use absolute or relative measures when assessing foodscape exposure in relation to fruit and vegetable intake? Evidence from a wide-scale Canadian study. *Prev Med* 2015;71:83-87.
32. 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(3):231-236.
33. Smith NR, Clark C, Fahy AE, et al. The Olympic Regeneration in East London (ORiEL) study: protocol for a prospective controlled quasi-experiment to evaluate the impact of urban regeneration on young people and their families. *BMJ Open* 2012;2(4):1-11.
34. Ordnance Survey, Southampton, UK. OS MasterMap® Address Layer and Address Layer 2: User Guide. Available at: http://www.ordnancesurvey.co.uk/oswebsite/docs/user-guides/os-mastermap-address-layer-and-al2-user-guide.pdf. Published January 2011. (Accessed).
35. Cummins S, Macintyre S. Are secondary data sources on the neighbourhood food environment accurate? Case-study in Glasgow, UK. *Prev Med* 2009;49(6):527-528.
36. Sallis JF, Nader PR, Rupp JW, et al. San Diego surveyed for heart-healthy foods and exercise facilities. *Public Health Rep* 1986;101(2):216-219.

37. Levine N. Crime mapping and the Crimestat program. *Geographical analysis* 2005;38(1):41-56.
38. Silverman BW. *Density estimation for statistics and data analysis*. New York, NY: Chapman and Hall; 1986.
39. Drewnowski A, Aggarwal A, Hurvitz PM, et al. Obesity and supermarket access: proximity or price? *Am J Public Health* 2012;102(8):74-80.
40. LeDoux TF, Vojnovic I. Going outside the neighborhood: The shopping patterns and adaptations of disadvantaged consumers living in the Lower eastside neighborhoods of Detroit, Michigan. *Health Place* 2013;19:1-14.
41. Kerr J, Frank L, Sallis JF, et al. Predictors of trips to food destinations. *Int J Behav Nutr Phys Act* 2012;9(58):1-10.
42. Hillier A, Cannuscio CC, Karpyn A, et al. How far do low-income parents travel to shop for food? Empirical evidence from two urban neighborhoods. *Urban Geogr* 2011;32(5):712-729.
43. Day RA. Consumer shopping behaviour in a planned urban environment. *Tijdschrift voor economische en sociale geografie* 1973;64(2):77-85.
44. Chum A, Farrell E, Vaivada T, et al. The effect of food environments on fruit and vegetable intake as modified by time spent at home: a cross-sectional study. *BMJ Open* 2015;5(6):e006200.
45. Lebel A, Kestens Y, Pampalon R, et al. Local context influence, activity space, and foodscape exposure in two canadian metropolitan settings: is daily mobility exposure associated with overweight? *J Obesity* 2012;2012(912645):1-9.
46. Kestens Y, Lebel A, Chaix B, et al. Association between Activity Space Exposure to Food Establishments and Individual Risk of Overweight. *PLoS ONE* 2012;7(8):e41418.
47. Wang MC, Kim S, Gonzalez AA, et al. Socioeconomic and food-related physical characteristics of the neighbourhood environment are associated with body mass index. *J Epidemiol Community Health* 2007;61(6):491-498.
48. Nakaya T. Geographically weighted generalised linear modeling. In: Brunsdon C, Singleton A, eds. *Geocomputation: A Practical Primer*. London, UK: SAGE, 2015:201-220.
49. Johnson JB, Omland KS. Model selection in ecology and evolution. *Trends Ecol Evol* 2004;19(2):101-108.
50. FSA Nutrient and Food Based Guidelines for UK Institutions. Food Standards Agency, 2006.
51. Longacre MR, Drake KM, MacKenzie TA, et al. Fast-food environments and family fast-food intake in nonmetropolitan areas. *Am J Prev Med* 2012;42(6):579-587.
52. Inagami S, Cohen DA, Brown AF, et al. Body mass index, neighborhood fast food and restaurant concentration, and car ownership. *J Urban Health* 2009;86(5):683-695.
53. Wooldridge J. Omitted Variable Bias: The Simple Case. *Introductory econometrics: A modern approach*: Cengage Learning, 2012.
54. Mehta NK, Chang VW. Weight status and restaurant availability a multilevel analysis. *Am J Prev Med* 2008;34(2):127-133.
55. Christian WJ. Using geospatial technologies to explore activity-based retail food environments. *Spatial and Spatio-temporal Epidemiology* 2012;3(4):287–295.
56. Babey SH, Wolstein J, Diamant AL. Food environments near home and school related to consumption of soda and fast food. *UCLA Center for Health Policy Research Policy Briefs* 2011;Jul 2011(PB2011-6):1-8.
57. Mercille G, Richard L, Gauvin L, et al. Associations between residential food environment and dietary patterns in urban-dwelling older adults: Results from the VoisiNuage study. *Public Health Nutr* 2012;15(11):2026–2039.

58. Van Hulst A, Barnett TA, Gauvin L, et al. Associations between children's diets and features of their residential and school neighbourhood food environments. *Can J Public Health* 2012;103(9 Suppl 3):48-54.
59. Lytle LA. Measuring the food environment: state of the science. *Am J Prev Med* 2009;36(4 Suppl):134-144.
60. Cummins S. Neighbourhood food environment and diet: time for improved conceptual models? *Prev Med* 2007;44(3):196-197.
61. Diez Roux AV. Investigating neighborhood and area effects on health. *Am J Public Health* 2001;91(11):1783-1789.
62. Health and Social Care Information Centre LS. "Statistics on Obesity, Physical Activity and Diet: England, 2013". 2013.
63. Li R, Serdula M, Bland S, et al. Trends in fruit and vegetable consumption among adults in 16 US states: Behavioral Risk Factor Surveillance System, 1990-1996. *Am J Public Health* 2000;90(5):777-781.
64. Shareck M, Frohlich KL, Kestens Y. Considering daily mobility for a more comprehensive understanding of contextual effects on social inequalities in health: a conceptual proposal. *Health Place* 2014;29:154-160.
65. Matthews SA, Yang TC. Spatial Polygamy and Contextual Exposures (SPACEs): Promoting Activity Space Approaches in Research on Place and Health. *Am Behav Sci* 2013;57(8):1057-1081.
66. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavorial medicine. *Ann Behav Med* 1994;16(3):199-202.
67. Wheeler D, Tiefelsdorf M. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *J Geogr Systems* 2005;7(2):161-187.

Tables

Table 1. Socio-demographic Characteristics and Fruit and Vegetable Intakes of Individuals (n=969), Baseline of the Olympic Regeneration in East London (ORiEL) Study, London, United-Kingdom, 2012

	No.	%	
Sociodemographic characteristics			
Age [a]	40.25	(8.21)	
Sex			
Female	732	75.5	
Male	237	24.5	
Ethnic origin			
White British or Irish	242	25.0	
Other White background	134	13.8	
Asian or Asian British	323	33.3	
Black or Black British	218	22.5	
Other ethnic background, including mixed background	52	5.4	
Highest qualification level			
None	220	22.7	
Low	283	29.2	
Intermediate	83	8.6	
High	188	19.4	
Foreign	195	20.1	
Marital status			
Married or in a civil partnership	618	63.8	
Separated, but still legally married or in a civil partnership; Divorced; Widowed	123	12.7	
Never married or never in a civil partnership	228	23.5	
Time-budget			
Low time-budget	193	19.9	
Intermediate time-budget	529	54.6	

High time-budget	206	21.3	
Non applicable	41	4.2	
Food intakes			
Daily fruit consumption			
0 portion	70	7.2	
1 portion	176	18.2	
2 portions	280	28.9	
3 portions	258	26.6	
4 portions	112	11.6	
5 or + portions	73	7.5	
Daily vegetable consumption			
0 portion	47	4.9	
1 portion	204	21.1	
2 portions	308	31.7	
3 portions	227	23.4	
4 portions	103	10.6	
5 or + portions	80	8.3	
Daily fruit and vegetable consumption			
< 5 portions	469	48.4	
5 or + portions	500	51.6	
Abbreviations : SD, standard deviation. [a] Value is expressed as mean (SD)			

Table 2. Measures of the Food Environment Exposure at Residential Address of ORIEL Participants (n=969), London, United-Kingdom, 2012

	Mean (SD)	Minimum	Maximum	Spearman r Value (p-value)
Absolute densities around home				
All healthy outlets (nb of healthy outlets/km²)	2.65 (2.26)	0.13	15.03	
All unhealthy outlets (nb of unhealthy outlets/km²)	16.33 (9.05)	1.26	50.66	
Relative density around home				
Percentage of healthy outlets (%)	12.55 (5.83)	3.32	37.42	
Correlation between densities of healthy and unhealthy outlets around home				0.866 (<0.001)
Abbreviations : SD, standard deviation.				

Table 3. Global and Local Modelling of the Relationships Between Different Food Environment Exposures and Fruit and Vegetable Intake for ORiEL Participants (n=969), London, United-Kingdom, 2012

Model	OLS				GWR			
	β	95% CI	AICc	Moran's *I*	β [g]	AICc	Selected Bandwidth Minimizing the AICc	Moran's *I*
1[a]	0.037	0.00, 0.07	3248.56	0.0258 [j]	-0.121, 0.14	3232.11	120	-0.0022
2[b]	0.003	- 0.01, 0.01	3252.09	0.0257 [j]	-0.098, 0.061	3220.73	48	-0.0073
3[c]	0.022 [i]	0.01, 0.04	3243.88	0.0239 [j]	0.009, 0.039	3232.65	269	0.006
4[d]	0.023	-0.01, 0.06	3196.53	0.0254 [j]	-0.584, 0.187	3163.85	70	-0.0075
5[e]	-0.003	-0.01, 0.01	3197.68	0.0251 [j]	-0.022, 0.019	3175.73	259	0.0038
6[f]	0.022 [i]	0.01, 0.04	3189.29	0.0242 [j]	0.01, 0.045	3176.39	253	0.0025

Abbreviations : AICc, Akaike information criterion corrected; CI, confidence interval; GWR, geographically weighted regression; OLS, ordinary least squares

[a] Fully adjusted linear regression model estimating the association between healthy outlets' density (nb/km²) and fruit intake
[b] Fully adjusted linear regression model estimating the association between unhealthy outlets' density (nb/km²) and fruit intake
[c] Fully adjusted linear regression model estimating the association between %healthy outlets' density (nb/km²) and fruit intake
[d] Fully adjusted linear regression model estimating the association between healthy outlets' density (nb/km²) and vegetable intake
[e] Fully adjusted linear regression model estimating the association between unhealthy outlets' density (nb/km²) and vegetable intake
[f] Fully adjusted linear regression model estimating the association between %healthy outlets' density (nb/km²) and vegetable intake
[g] Range of locally estimated coefficients
[i] $P < 0.01$
[j] $P < 0.001$

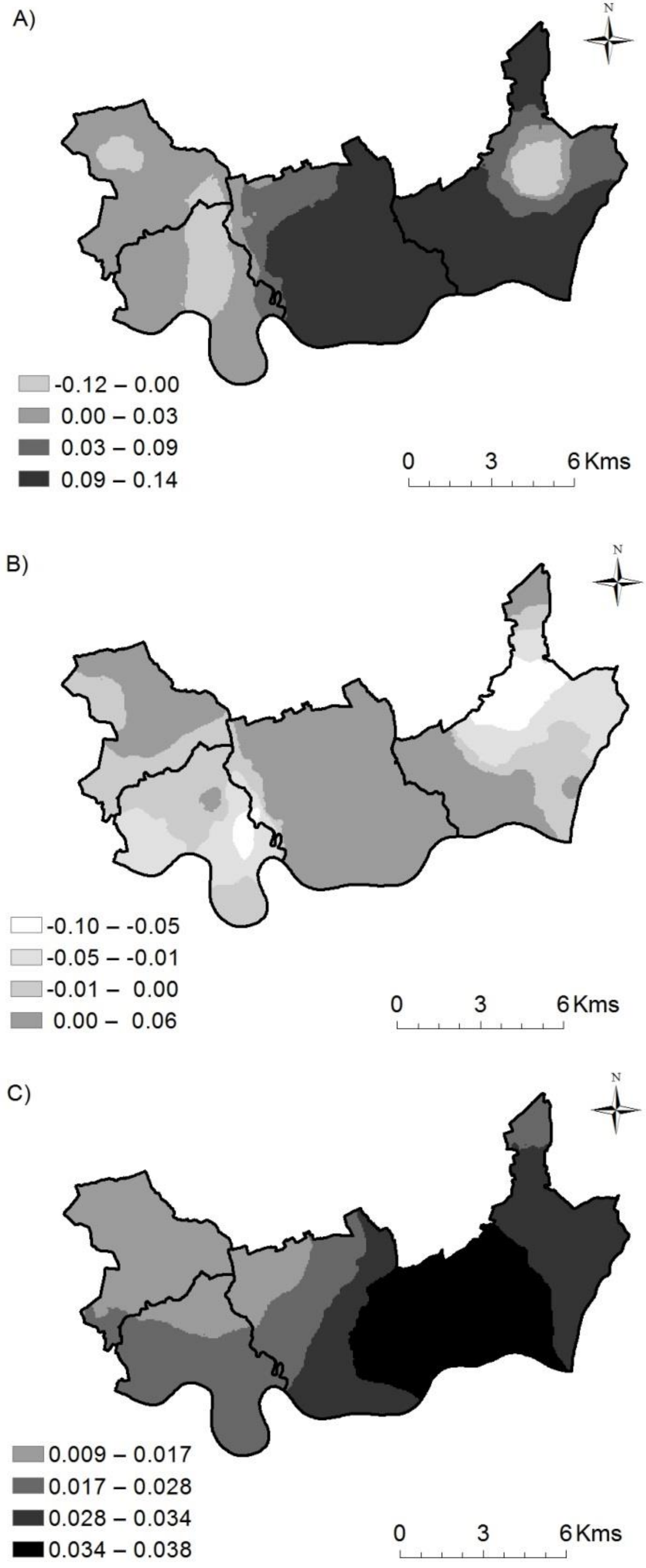
List of figures

Figure 1. Spatial variations of the estimated coefficients of the relation between three types of food environment exposures – A) density of healthy food outlets, B) density of unhealthy food outlets, and C) percentage of healthy outlets – and fruit intake for ORiEL participants (n=969), London, United-Kingdom, 2012. Cut-points of β-coefficients approximate quartiles, yet slightly modified so that positive and negative values of estimates are not mixed within a same colour hue.

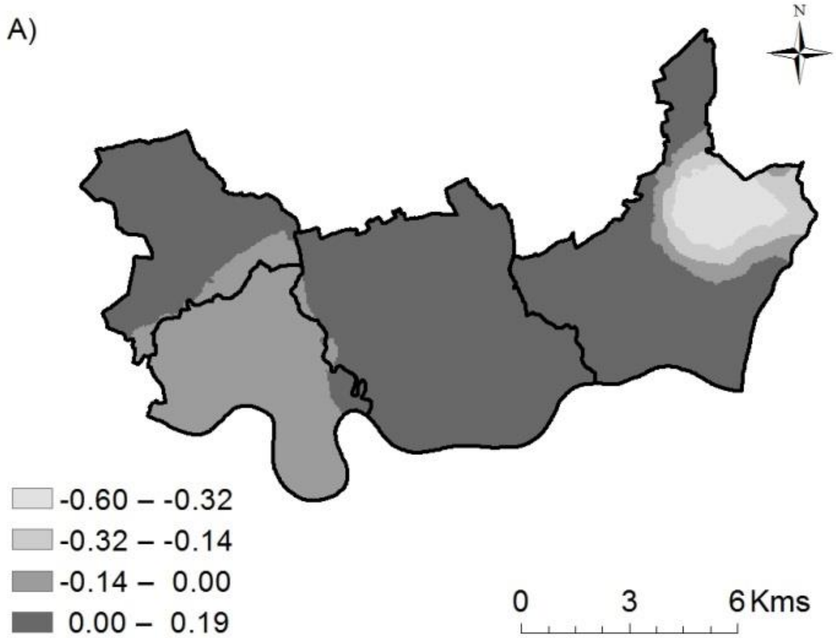
Figure 2 Spatial variations of the β-coefficients of the relation between three types of food environment exposures – A) density of healthy food outlets, B) density of unhealthy food outlets, and C) percentage of healthy outlets – and vegetable intake for ORiEL participants (n=969), London, United-Kingdom, 2012. Cut-points of β-coefficients approximate quartiles, yet slightly modified so that positive and negative values of estimates are not mixed within a same colour range.

Figure 3 Spatial variations of the t-values (absolute value) of the relation between three types of food environment exposures – A) density of healthy food outlets, B) density of unhealthy food outlets, and C) percentage of healthy outlets – and fruit intake, for ORiEL participants (n=969), London, United-Kingdom, 2012. A |t-value| ≥ 1.96 shows significant associations (P-value ≤ 0.05).

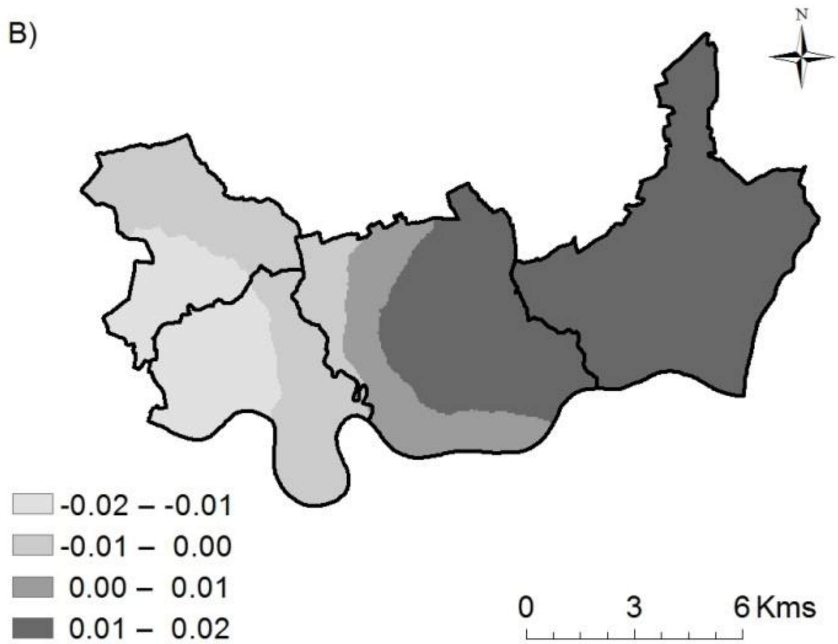
Figure 4 Spatial variations of the t-values (absolute value) of the relation between three types of food environment exposures – A) density of healthy food outlets, B) density of unhealthy food outlets, and C) percentage of healthy outlets – and vegetable intake, for ORiEL participants (n=969), London, United-Kingdom, 2012. A |t-value| ≥ 1.96 shows significant associations (P-value ≤ 0.05).


A)
N
-0.12 – 0.00
0.00 – 0.03
0.03 – 0.09
0.09 – 0.14
0 3 6 Kms
B)
N
-0.10 – -0.05
-0.05 – -0.01
-0.01 – 0.00
0.00 – 0.06
0 3 6 Kms
C)
N
0.009 – 0.017
0.017 – 0.028
0.028 – 0.034
0.034 – 0.038
0 3 6 Kms

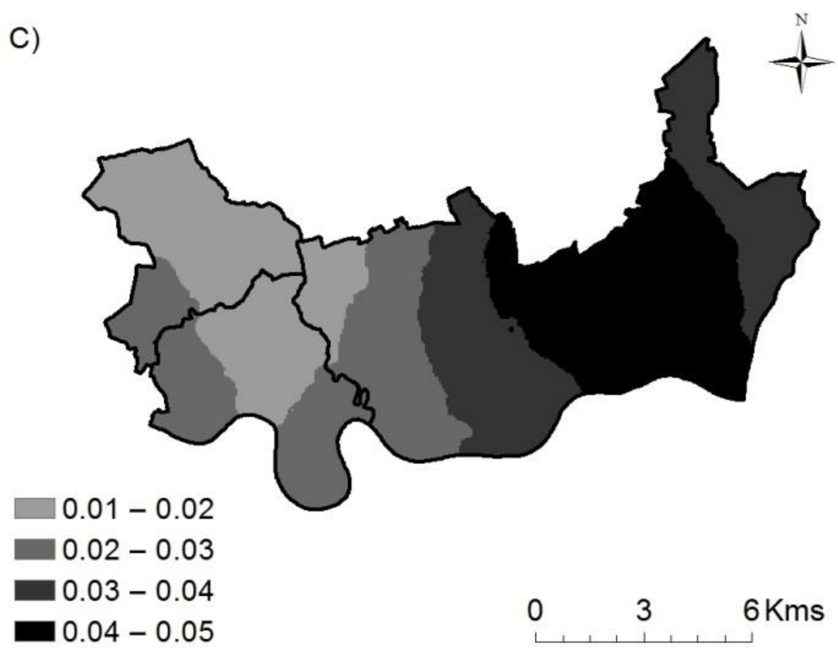
A)

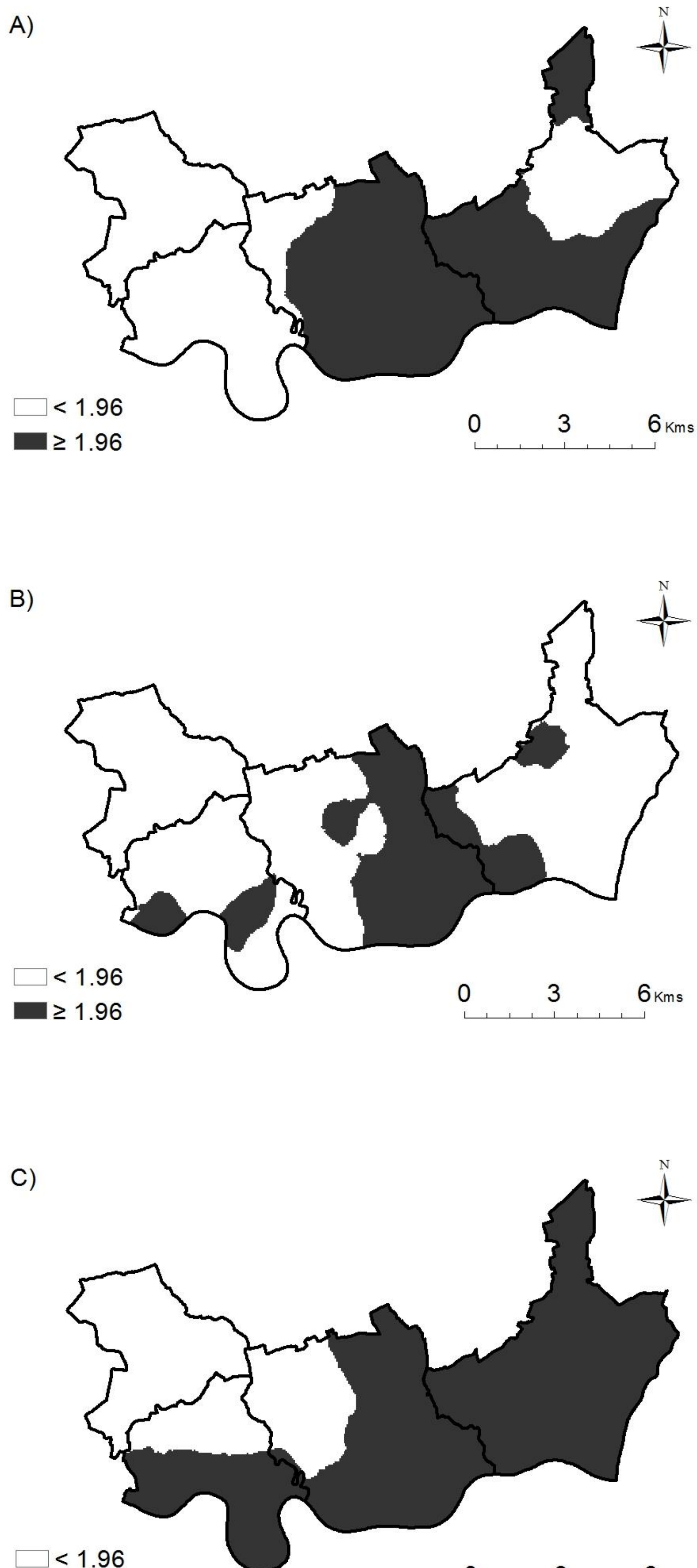

N
-0.60 – -0.32
-0.32 – -0.14
-0.14 – 0.00
0.00 – 0.19
0 3 6 Kms

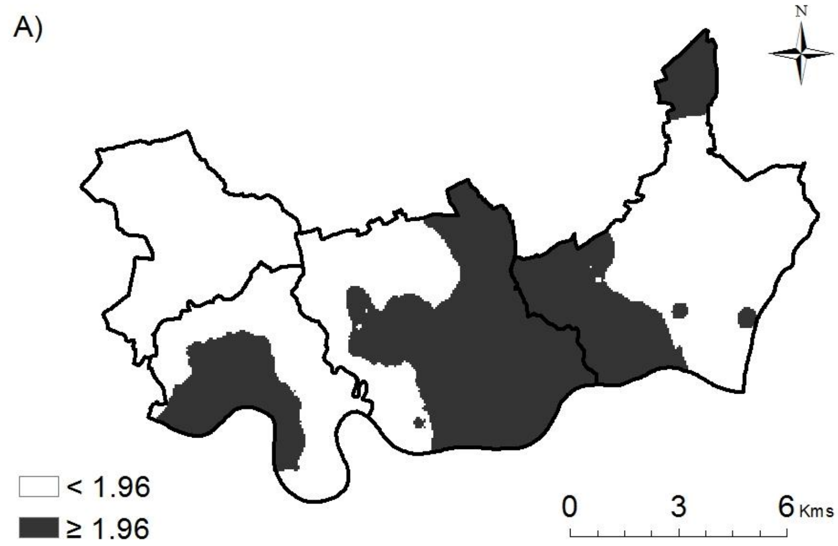
B)

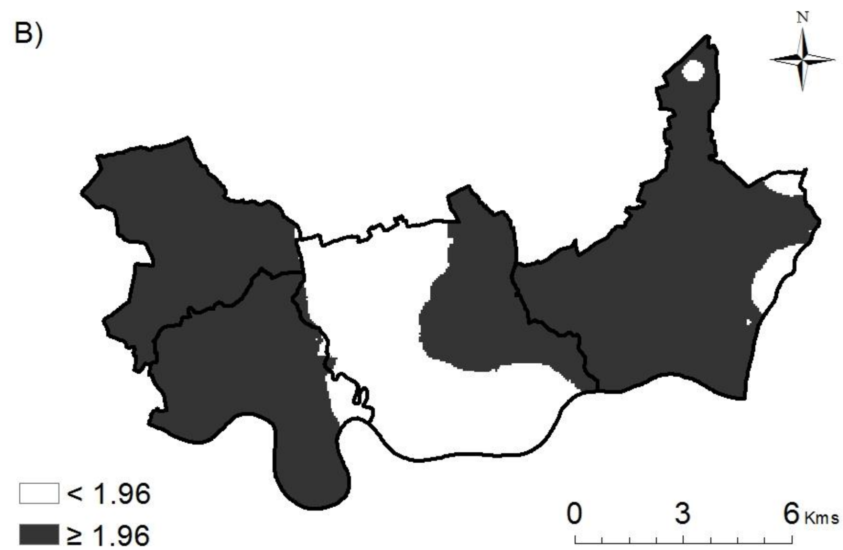

N
-0.02 – -0.01
-0.01 – 0.00
0.00 – 0.01
0.01 – 0.02
0 3 6 Kms

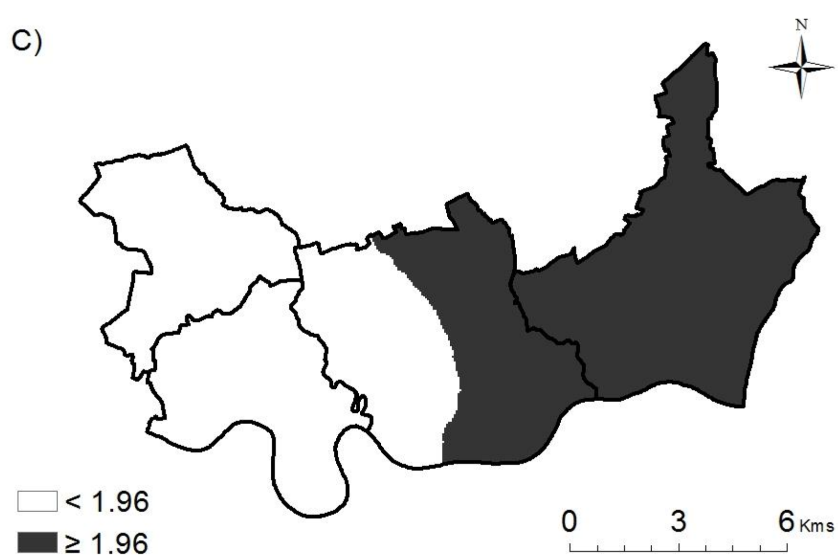
C)

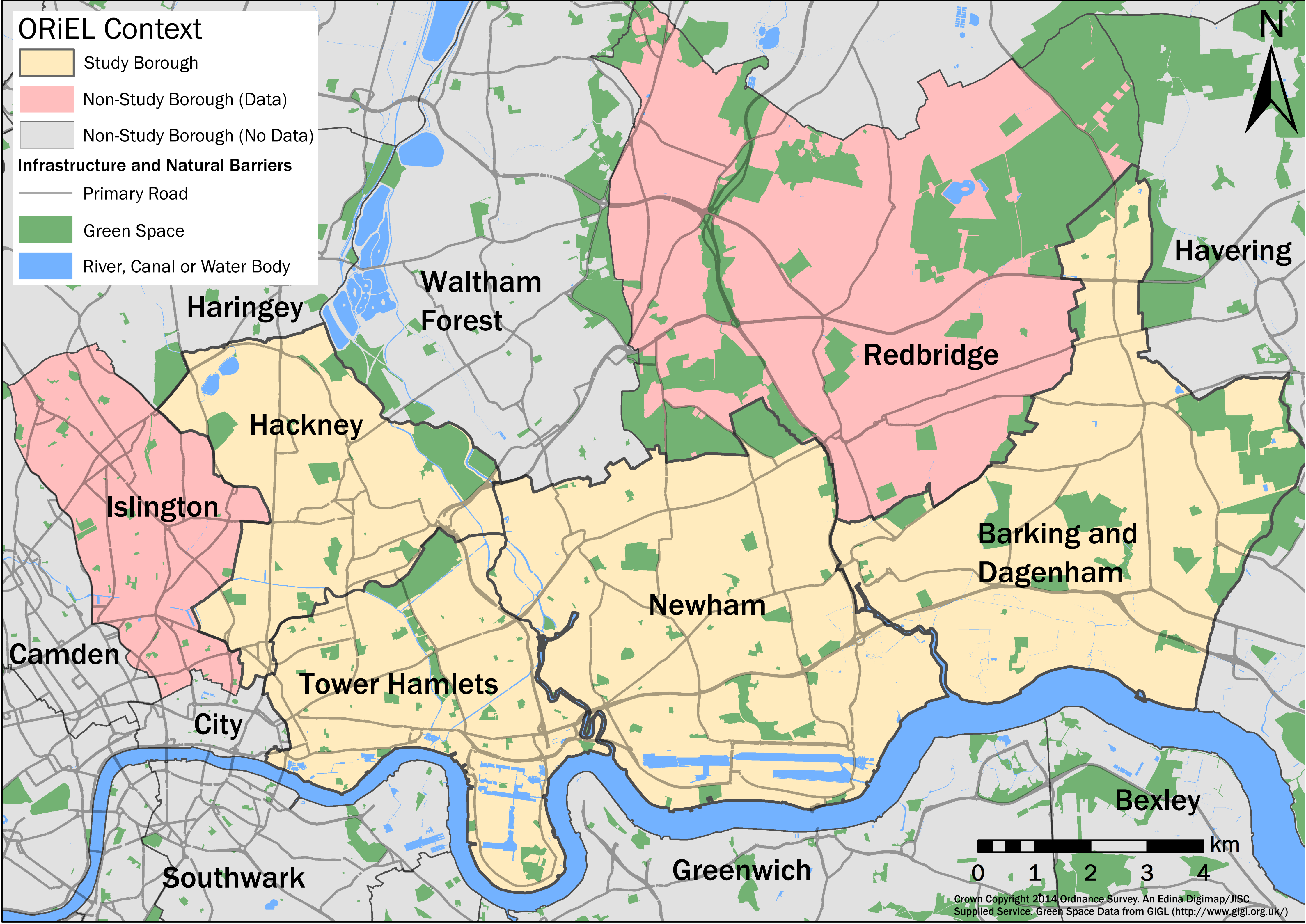

N
0.01 – 0.02
0.02 – 0.03
0.03 – 0.04
0.04 – 0.05
0 3 6 Kms


A)
N
< 1.96
≥ 1.96
0
3
6 Kms
B)
N
< 1.96
≥ 1.96
0
3
6 Kms
C)
N
< 1.96
≥ 1.96
0
3
6 Kms


A)
N
< 1.96
≥ 1.96
0
3
6
Kms


B)
N
< 1.96
≥ 1.96
0
3
6
Kms


C)
N
< 1.96
≥ 1.96
0
3
6
Kms



ORiEL Context
Study Borough
Non-Study Borough (Data)
Non-Study Borough (No Data)
Infrastructure and Natural Barriers
Primary Road
Green Space
River, Canal or Water Body
N
Haringey
Waltham Forest
Redbridge
Havering
Hackney
Islington
Barking and Dagenham
Newham
Camden
Tower Hamlets
City
Bexley
Southwark
Greenwich
0 1 2 3 4 km
Crown Copyright 2014 Ordnance Survey. An Edina Digimap/JISC Supplied Service. Green Space Data from GIGL (http://www.gigl.org.uk/)

Web Table 1. Descriptive Data of the Food Environment of Six East-London Boroughs[a] (Newham, Tower Hamlets, Hackney, Barking & Dagenham, Islington and Redbridge), London, United-Kingdom, 2012

Type of Food Outlets	Description	Nutritional Class	No. of Food Outlets	Kernel density estimation Bandwidth size: 7% nearest neighbors		
				Number of neighbors	*Distance(meters)* [b]	
					Mean	*SD*
Food Stores						
Chain Supermarkets	Nationally recognisable multi-store companies that are able to leverage supply to sell a wide range of products competitively (i.e. Tesco, Sainsburies, Asda, Morrisons).	Healthy	94	7	2414	1035
Independent Supermarkets	Generic non-chain supermarket.	Healthy	78	5	2054	961
Ethnic-Specific Supermarkets.	Independent supermarket that specialises in selling culturally/ethnically specific “world” food.	Healthy	136	10	2196	1623
Fruit and Vegetables stores	Green grocers, fruiterers.	Healthy	42	3	1940	878
Convenience Stores	Small store (i.e. corner shop, petrol station forecourt) selling a limited range of foods.	Unhealthy	1237	87	1960	617
Discount Retailers	Stores, either chain or independent, specifically dealing in discount foods (i.e. Lidl, Aldi, Iceland)	ND	72	.	.	.
“Pound Store” Retailers	General discount stores which sell a range of non-food items as well as long-life or dried food goods.	ND	57	.	.	.
Affiliated Food stores	Symbol group/franchise store (i.e. Budgens, Spar, Costcutter, Nisa).	ND	167	.	.	.
Specialist Food Stores	Food store focusing on particular niches: butchers, fishmongers, health foods, bakers, confectioners etc.	ND	307	.	.	.
Food Services						
Fast-Food Restaurants (Chain&Franchised)	A multi-premises restaurant business that offers food and drink in a self-service manner to eat in, or by collection or delivery to take away.	Unhealthy	86	6	2031	667
Fast-Food Restaurants (Independent)	As above, but for independent restaurants.	Unhealthy	1064	74	1895	579
Full Service Restaurants	A restaurant offering a selection of foods and beverages in addition to table service.	ND	777	.	.	.
Cafes, Coffee Shops and Sandwich bars	Chain and non-chain sandwich, snack and coffee bars (i.e. Subway, Starbucks, Greggs, Percy Ingle)	ND	1037	.	.	.
Pubs and Bars	A drinking establishment that also provides meals.	ND	671	.	.	.
Other						
Entertainment or sport focussed food retailers, and private food businesses	Cinemas, theatres, leisure activities, sports clubs, sports centres and other sporting venues that also sell food. Medical, schools, caring establishments, catering, wholesalers (where membership is required), and light food industry.	ND	439	.	.	.

Abbreviations : ND, non-determined, SD, standard deviation.

[a] Notes on the validation of the food environment dataset. A 10% stratified random sample (n= 604) of the list of registered food businesses at the time of ORiEL study baseline (i.e. 2012) was taken from the four main ORiEL study boroughs (Hackney, Tower Hamlets, Newham, Barking & Dagenham). Validation was performed based on the match between each business in the sample and age concomitant Google Streetview street photographies. Matches (true positives) and mismatches (false positives) were assigned where the business was respectively the same as, and different from, the photographed businesses. False negatives were not assessed, because we did not have a dataset of food businesses that exist but are not recorded by the registration data. However, as it is a mandatory requirement under UK food safety legislation for businesses to be registered, it is unlikely that food businesses would be operating without having been recorded by the local authority. False negatives are therefore likely to be infrequent. The positive predictive value (PPV) for the food businesses was 0.96 (0.94, 0.98 - Clopper-Pearson exact binomial CI). This implies that the dataset contains 96% valid records. Chi-squared tests showed no significant differences by borough, deprivation (tertiles of Index of Multiple Deprivation), or type of food business (food stores versus restaurants). For the reasons evoked above, sensitivity could not be assessed, but is expected to be high.

[b] Measures of central tendency (mean) or spread (standard deviation) for the radius of the set of circular home-centered buffers encapsulating 7% of the point distribution of the food outlet type

Web Table 2. Impact of Adaptive Bandwidth Size on the Associations Between Food Environment Exposure and Fruit and Vegetable Intakes for ORiEL Participants (n=969), London, United-Kingdom, 2012

		Outcome - Fruit intake						Outcome - Vegetable intake					
		Model 1		**Model 2**		**Model 3**		**Model 4**		**Model 5**		**Model 6**	
Adaptive bandwidth size	**Foodscape exposures**	β	*p-value*	β	*p-value*	β	*p-value*	β	*p-value*	β	*p-value*	β	*p-value*
1%	Healthy outlets	.005	.539					-.001	.856				
	Unhealthy outlets			.002	.381					.000	.871		
	%healthy outlets					.000	.948					.001	.887
3%	Healthy outlets	.016	.087					.003	.716				
	Unhealthy outlets			.005	.267					-.001	.822		
	% healthy outlets					.007	.182					.010	.062
5%	Healthy outlets	.038	.020					.023	.146				
	Unhealthy outlets			.004	.430					-.003	.511		
	% healthy outlets					.017	.008					.017	.005
7%	Healthy outlets	.037	.052					.023	.204				
	Unhealthy outlets			.003	.568					-.003	.489		
	% healthy outlets					.022	.004					.022	.003
10%	Healthy outlets	.052	.018					.032	.135				
	Unhealthy outlets			.003	.546					-.002	.660		
	% healthy outlets					.012	.011					.011	.015