doi: 10.1111/obr.12378

Supplement Article

Neighbourhood typology based on virtual audit of environmental obesogenic characteristics

T. Feuillet,[1] H. Charreire,[1,2] C. Roda,[1] M. Ben Rebah,[1] J. D. Mackenbach,[3] S. Compernolle,[4] K. Glonti,[5] H. Bárdos,[6] H. Rutter,[5] I. De Bourdeaudhuij,[4] M. McKee,[5] J. Brug,[3] J. Lakerveld[3] and J.-M. Oppert[1,7]

[1]Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Centre de Recherche en Epidémiologie et Statistiques, Cnam, COMUE Sorbonne Paris Cité, Université Paris 13, Bobigny, France, [2]Paris Est University, Lab-Urba, UPEC, Urban School of Paris, Créteil, France, [3]Department of Epidemiology and Biostatistics, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands, [4]Department of Movement and Sport Sciences, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium, [5]ECOHOST – The Centre for Health and Social Change, London School of Hygiene and Tropical Medicine, London, UK, [6]Department of Preventive Medicine, Faculty of Public Health, University of Debrecen, Debrecen, Hungary, and [7]Sorbonne Universités, Université Pierre et Marie Curie, Université Paris 06, Institute of Cardiometabolism and Nutrition, Department of Nutrition, Pitié-Salpêtrière Hospital, Assistance Publique-Hôpitaux de Paris, Paris, France

Received 14 December 2015; accepted 15 December 2015

Address for correspondence: T. Feuillet, Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Centre de Recherche en Epidémiologie et Statistiques, Inserm (U1153), Inra (U1125), Cnam, COMUE Sorbonne Paris Cité, Université Paris 13, Bobigny, France. E-mail: thierry.feuillet@univ-paris8.fr

The copyright line for this article was changed on 30 May 2016 after original online publication

Summary

Virtual audit (using tools such as Google Street View) can help assess multiple characteristics of the physical environment. This exposure assessment can then be associated with health outcomes such as obesity. Strengths of virtual audit include collection of large amount of data, from various geographical contexts, following standard protocols. Using data from a virtual audit of obesity-related features carried out in five urban European regions, the current study aimed to (i) describe this international virtual audit dataset and (ii) identify neighbourhood patterns that can synthesize the complexity of such data and compare patterns across regions. Data were obtained from 4,486 street segments across urban regions in Belgium, France, Hungary, the Netherlands and the UK. We used multiple factor analysis and hierarchical clustering on principal components to build a typology of neighbourhoods and to identify similar/dissimilar neighbourhoods, regardless of region. Four neighbourhood clusters emerged, which differed in terms of food environment, recreational facilities and active mobility features, i.e. the three indicators derived from factor analysis. Clusters were unequally distributed across urban regions. Neighbourhoods mostly characterized by a high level of outdoor recreational facilities were predominantly located in Greater London, whereas neighbourhoods characterized by high urban density and large amounts of food outlets were mostly located in Paris. Neighbourhoods in the Randstad conurbation, Ghent and Budapest appeared to be very similar, characterized by relatively lower residential densities, greener areas and a very low percentage of streets offering food and recreational facility items. These results provide multidimensional constructs of obesogenic characteristics that may help target at-risk neighbourhoods more effectively than isolated features.

Keywords: Cluster analysis, SPOTLIGHT, obesogenic environment, virtual audit.

Abbreviations: GSV – Google Street View; MFA– multiple factor analysis; SES– socio economic status

obesity reviews (2016) **17** (Suppl. 1), 19–30

Background

Social ecological models of health behaviour (1) view the built environment as a key contextual determinant of various health outcomes, including overweight and obesity. The characteristics of neighbourhoods can indeed influence obesity-related behaviours such as dietary intake and physical activity (2–4). A better understanding of obesogenic features of the environment, which can help to identify distinct or contrasting types of neighbourhoods, is thus an important step in establishing policies to tackle the obesity epidemic (5–7).

Evidence on associations between built environment and obesity remains inconsistent, especially when comparing different countries (8). This may reflect two important issues, which are (i) the need for identifying instruments or measures that would help assess in a standardized way environmental features potentially related to obesity and (ii) the need for better means of capturing spatial interactions between these features, i.e. the specific relations between close environmental features compared with remote ones. The first issue poses a challenge to data collection. To delineate pathways through which the built environment influences obesity-related behaviours, the first step is to assess a comprehensive set of environmental features using standard protocols, in order to lessen the potential pitfalls of methodological variation between studies (9). Moreover, most studies of obesity and the built environment measured physical activity facilities or food outlets only (10–14). Yet, 'a comprehensive picture of the relation between the built environment and obesity' [14, p. 139] requires concurrent information on the broader spectrum of features of physical activity and food environments (15) using standard protocols to achieve comparability in the data collected and thus overcome inconsistencies due to measurement techniques (9). Remote sensing tools, such as Google Street View (GSV), offer innovative methods to assess built environment characteristics in different geographical contexts (16–21). The first objective of this study was thus to apply such a GSV-based tool, already shown as valid and reliable by Bethlehem *et al.* (22), in different neighbourhoods in different European countries, in order to achieve an environmental dataset that would allow for future comparisons of epidemiological analyses between countries. We were interested to know if this descriptive approach might reveal some meaningful similarities or differences between neighbourhoods of different countries, which would not have been possible without such a tool. The second issue relates to how the complexity of such built environmental data should be synthesized. Obesity-related environmental features (e.g. cycle paths, fast-food restaurants, etc.) interact spatially, so that the association of one pair of features can be moderated by the presence of others. Indeed, obesogenic environments 'are characterized by clustered factors that promote excess caloric intakes and inhibit physical activity' (4). Because these characteristics are not mutually exclusive and may be present at the same places, epidemiological models should consider multidimensional factors accounting for complexity of neighbourhood structure.

To address these challenges, we used a two-step approach based on factor analysis and hierarchical clustering. We used data collected in the European Commission-funded Sustainable Prevention of Obesity through Integrated Strategies (SPOTLIGHT) project (23,24). A GSV-based virtual audit tool was developed to assess potential obesogenic environmental characteristics (22) and used in 59 neighbourhoods in five European urban regions.

The objectives were threefold:

1. To describe the data collected with the GSV virtual audit tool on the food and physical activity environment and to explore potential heterogeneity in the five European urban regions;
2. To build a typology of neighbourhoods based on these obesogenic environmental characteristics, using a factor analysis-based approach;
3. To explore whether neighbourhood types were country-specific.

Methods

Study design and neighbourhood sampling

Twelve neighbourhoods were selected in each of the five included urban regions across Europe: Ghent and suburbs (Belgium), Paris and inner suburbs (France), Budapest and suburbs (Hungary), the Randstad (including Amsterdam, Rotterdam, The Hague and Utrecht in the Netherlands) and Greater London (UK). All these neighbourhoods were based on small-scale local administrative neighbourhoods as used in each country, except for Hungary, because Budapest is divided into districts and suburbs that are highly heterogeneous in terms of population and surface area. In order to ensure comparability between study areas, we thus defined 1 km^2 areas to represent neighbourhoods in Budapest and suburbs. As described elsewhere (24), the neighbourhood sampling sought to include a mix of environmental contexts based on residential density and neighbourhood socioeconomic status (SES). Data on residential density were obtained from the Urban Atlas database (European Environment Agency, 2002). Because this database is 14 years old, it is likely that residential density has changed since then; however, the Urban Atlas is a unique resource providing a comparable index of land use across European countries. Among six existing residential density categories in the Urban Atlas, only two classes were used: high and low residential density (corresponding to >80% and <50% of areas covered by residential buildings, respectively). SES levels were derived from the most recent income data national censuses available (dated between 2008 and 2010), with two classes:

low and high (i.e. the first and third tertiles). This combination resulted in four neighbourhood types (high residential density/high SES, high residential density/low SES, low residential density/high SES and low residential density/low SES). In addition, sampled neighbourhoods had to contain a minimum threshold of adult inhabitants. For a target sample of about 100 residents in each neighbourhood, with an estimate of 10% response rate, approximately 1,000 residents were sampled in each neighbourhood. Given varying expected response rate according to neighbourhood SES (25), 1,200 adults were sampled in low SES neighbourhoods and 800 in high SES neighbourhoods. Finally, a total of 12 neighbourhoods per neighbourhood type were selected. Hence, three neighbourhoods from each category were randomly selected in each region, leading to a total of 60 neighbourhoods.

The Google Street View-based virtual audit tool and database

As previously described (22), the SPOTLIGHT virtual audit tool was designed for the assessment of food and physical activity-related features of the built environment within neighbourhoods. It initially contained 42 items, grouped into eight domains: walking (six items), cycling (eight items), public transport (two items), aesthetics (nine items), land use mix (three items), grocery stores (five items), food outlets (six items) and recreational facility-related items (three items). We removed items that appeared twice in different categories, resulting in 36 items (Table 1). These items were assessed in all streets of 59 neighbourhoods (online release of GSV in Hungary only happened late in the study resulting in the fact that one Hungarian neighbourhood that had been sampled was actually not assessed by GSV), giving 4,486 street segments. The virtual audit was carried out in 2014 in each country by trained researchers of the SPOTLIGHT project team. The virtual audit of one neighbourhood took approximately one week, on average. The validity (vs. field audit) and reliability (using test-retest) of the tool was previously demonstrated in a random sample of streets from four Dutch neighbourhoods (22).

All built environmental features were then aggregated from street segment to the neighbourhood level by taking the percentage of street segments in each neighbourhood that contained them. For instance, the item 'type of street' includes three features: pedestrian friendly street, regular road and road with high-speed traffic, so this item generated three discrete variables at the neighbourhood level. If 100 of the 500 street segments of a neighbourhood were qualified as 'pedestrian friendly', then the feature 'pedestrian friendly street' was quantified as 0.20 in this neighbourhood. Finally, a total of 56 environmental features were assessed (Table 1). A detailed description of each has been previously provided by Bethlehem *et al.* (22). Regions and types of neighbourhoods (four categories defined by residential density and SES level) were added as illustrative categorical variables in the factor analysis.

Factor and cluster analyses

Exploratory factor analysis reveals structure and pattern in a data matrix (26). This method has been applied before in studies of neighbourhood patterns related to health (11,12,27–29). To account for the existing structure of the data – i.e. the eight domains of the GSV data – we performed a multiple factor analysis (MFA). MFA is a subset of principal component analysis (which is, in turn, a subset of exploratory factor analysis) in which the variables (i.e. items from virtual audit) in the same domain are weighted to balance the importance of the domains. This weighting is a function of the highest axial inertia of each domain (30). MFA leads to identification of dimensions, as in principal component analysis, which forms the basis for neighbourhood clustering through a hierarchical clustering on principal components (Ward's method (31)).

In a first step, MFA was used to determine the dimensions that synthesized most information. The number of dimensions retained was determined as eigenvalue (i.e. the variance in all the variables that is accounted for by that factor) >1, jump in the scree plot (plot of the total variance related to each dimension) and interpretability of MFA (12,26). To facilitate interpretation, associations between initial variables and dimensions retained were assessed, using Pearson correlation for continuous variables (i.e. the 56 categories), or one-way analysis of variance for categorical variables (European urban regions and type of neighbourhoods). The Fisher test was used to assess whether the categorical variable was associated with the dimension, and t-tests were conducted category by category to determine whether the coordinates of individuals in the sub-population defined by each category were significantly different from the overall (i.e. different from 0) (32). In a second step, hierarchical clustering on principal components with k-means consolidation (i.e. a number of iterations of k-means algorithm were applied to the partitioning derived from the hierarchical clustering, for consolidating the clusters) was performed on the dimensions retained from the previous step, to identify clusters. The number of clusters was determined by maximizing variance between groups and minimizing variance within groups. Those clusters were characterized and interpreted with v-tests, relating each original variable to clusters. V-tests are based on the difference between the mean of the variable in the cluster and the overall mean of the variable. Higher absolute values indicate better characterization of the cluster (26). These values thus enable classifying categories of the variable by order of importance, facilitating interpretation of results (12). Because v-test values follow Student's

Table 1 Descriptive statistics of the Google Street View-based data of the SPOTLIGHT project (8 categories, 36 items, 56 modalities), by country and by type of neighbourhoods. All values are given in percentages

Variables	Total (%)	Countries (all values are in %)					Type of neighbourhoods*(all values are in %)			
		Ghent and suburbs (Belgium)	Paris and inner suburbs (France)	Budapest and suburbs (Hungary)	The Randstad (The Netherlands)	Greater London (UK)	High residential density – High SES	High residential density – Low SES	Low residential density – High SES	Low residential density – Low SES
Walking										
Type of street										
Pedestrian friendly street	2.2	2.4	0.8	0.5	6.8	0.5	1.4	3.7	0.1	4.0
Regular road	95.9	96.2	99.2	99.3	91.3	93.7	97.8	96.1	94.9	95.2
Road with high-speed traffic	1.9	1.4	0.0	0.2	1.9	5.8	0.8	0.2	5.0	0.8
Sidewalk										
Good	37.9	27.2	60.4	14.6	68.9	43.0	49.8	43.7	35.0	25.0
Fair	35.4	17.3	25.8	60.9	5.9	40.0	25.1	44.8	34.8	37.1
Poor	5.3	3.6	6.1	7.7	0.2	7.0	3.1	2.9	4.7	10.3
No sidewalk	21.0	51.9	4.7	16.7	24.7	9.6	21.5	8.3	25.0	27.2
Under construction	0.4	0.0	3.0	0.1	0.3	0.4	0.5	0.2	0.6	0.4
Pedestrian crossing										
No crossing	79.9	72.1	21.3	89.3	89.4	83.2	72.0	75.9	84.6	84.8
Traffic lights	5.7	4.1	6.6	4.3	5.9	8.3	11.0	8.5	2.3	2.6
Zebra-path	23.0	23.9	81.2	6.4	4.6	8.6	17.1	15.6	13.2	12.6
Street light										
Yes	96.0	93.0	90.6	99.7	97.7	92.8	96.6	98.7	91.8	98.4
Cycling										
Traffic calming devices										
Yes	26.7	23.6	47.4	4.9	59.1	24.0	28.7	33.0	22.3	24.7
Speed limit										
≤30	48.9	20.1	24.8	30.8	74.2	80.0	56.4	52.4	46.0	42.7
50	48.3	75.7	75.2	69.0	23.0	13.4	42.8	47.3	48.2	54.4
>50	2.8	4.2	0.0	0.2	2.8	6.6	0.8	0.3	5.8	2.9
Obstacles present on bicycle lanes										
No obstacle	13.8	18.4	19.6	3.9	17.4	9.7	15.6	13.5	7.2	11.2
No bicycle lane	85.7	80.9	80.4	96.1	82.0	89.0	83.8	85.3	92.5	88.5
Temporary	0.6	0.7	0.0	0.0	0.6	1.4	0.6	1.2	0.3	0.3
Cars form an obstacle on the road										
Yes	33.6	35.3	9.4	46.7	31.7	22.8	30.5	42.4	37.2	24.3
Public bicycle facilities										
Yes	1.2	0.1	1.4	0.0	0.5	4.3	2.6	2.5	0.2	0.1
Type of bicycle lanes										
On road cycle lane with markings	6.5	7.1	17.1	3.0	6.2	7.9	10.0	6.9	3.7	6.7
Shared path with pedestrians	0.8	1.2	2.5	0.4	0.4	1.1	0.9	0.5	1.3	0.5
Separate cycle lane with buffer	4.6	10.4	0.0	0.5	11.4	2.5	5.7	7.3	2.5	4.2
Public transport										
Bus/tram stop or shelter										

(Continues)

Table 1 (Continued)

Variables	Total (%)	Ghent and suburbs (Belgium)	Paris and inner suburbs (France)	Budapest and suburbs (Hungary)	The Randstad (The Netherlands)	Greater London (UK)	High residential density – High SES	High residential density – Low SES	Low residential density – High SES	Low residential density – Low SES
		Countries (all values are in %)					Type of neighbourhoods*(all values are in %)			
Yes	12.6	11.1	16.0	11.5	7.9	18.5	15.4	13.7	11.0	11.3
Railway/underground station										
Yes	0.6	0.0	1.4	0.3	0.3	1.3	1.1	0.5	0.4	0.4
Aesthetics										
Green and water areas visible										
Yes	33.2	49.4	11.0	7.1	55.4	48.8	43.6	19.7	30.7	39.5
Residential gardens										
Yes	66.2	66.8	52.1	72.4	59.6	67.8	56.3	41.2	75.8	86.2
Trees										
Yes	87.9	86.7	52.9	96.0	92.7	84.7	80.8	83.8	91.1	94.0
Litter										
Yes	17.6	9.0	12.9	3.8	41.0	24.0	12.8	25.2	16.0	17.3
Graffiti										
Yes	9.5	3.3	10.7	11.7	18.5	1.9	7.6	21.7	7.7	2.4
Rating of condition for most residential buildings										
Poor condition	0.3	0.2	0.3	0.3	0.0	0.6	0.1	0.8	0.3	0.0
Fair condition	25.9	3.5	19.3	55.5	1.0	21.7	18.3	34.3	26.7	24.2
Well-kept condition	64.6	86.2	71.3	39.4	88.5	63.7	70.8	56.5	63.0	68.5
Abandoned building or vacant area										
Yes	3.8	0.1	8.0	0.6	7.1	6.5	5.1	6.1	3.3	1.2
Land use mix										
Residential buildings visible										
Yes	90.9	90.0	91.2	95.3	89.5	86.0	89.2	91.6	90.0	92.7
Type of residential buildings										
Apartment above shops	4.0	3.2	7.4	5.2	2.9	2.6	10.0	4.4	2.0	0.8
Apartment buildings <5 stories	19.9	5.3	51.5	16.5	19.2	23.5	27.3	40.1	8.9	8.6
Apartment buildings >5 stories	3.5	1.1	7.7	2.7	2.6	5.6	4.1	8.0	1.0	2.1
Detached/semidetached homes	50.4	60.2	19.8	68.3	27.4	49.6	29.7	24.7	69.1	69.1
Terraced homes	13.0	20.0	4.7	2.6	37.4	4.7	18.1	14.4	9.0	12.2
Percentage of non-residential buildings compared with residential buildings										
Non-residential buildings	9.4	3.9	13.6	9.7	6.0	13.7	12.3	13.7	5.6	5.7
Grocery stores										
Supermarket										
Yes	2.3	2.8	4.4	1.5	1.6	3.0	2.7	3.2	1.5	2.1
Local shop										
Yes	3.1	2.7	8.8	3.0	2.4	2.2	4.9	5.0	1.3	2.1
Street food market										

(Continues)

Table 1 (Continued)

Variables	Total (%)	Countries (all values are in %) Ghent and suburbs (Belgium)	Paris and inner suburbs (France)	Budapest and suburbs (Hungary)	The Randstad (The Netherlands)	Greater London (UK)	Type of neighbourhoods*(all values are in %) High residential density – High SES	High residential density – Low SES	Low residential density – High SES	Low residential density – Low SES
Yes	0.1	0.0	0.3	0.0	0.1	0.3	0.1	0.2	0.1	0.1
Wine/liquor store										
Yes	0.4	0.0	0.6	0.4	0.5	0.3	0.5	0.9	0.1	0.1
Convenience store/small grocery store										
Yes	4.3	2.2	3.9	4.8	1.2	8.0	5.9	6.8	1.8	3.9
Food outlets										
On-street vendors of food										
Yes	0.3	0.0	0.3	0.3	0.5	0.3	0.2	1.0	0.1	0.0
Restaurant										
Yes	5.7	2.8	7.2	4.3	3.7	10.9	12.5	6.2	2.8	2.5
Fast food restaurant										
Yes	1.6	2.8	4.7	0.2	1.4	2.0	2.0	3.2	0.8	0.8
Take away restaurant										
Yes	1.8	0.0	0.8	1.2	0.6	5.3	3.4	2.5	0.4	1.5
Café/bar										
Yes	5.3	2.5	8.5	5.7	2.7	7.8	10.9	7.6	1.8	2.7
Shopping mall										
Yes	0.1	0.0	0.0	0.0	0.1	0.2	0.1	0.0	0.1	0.1
Physical activity facilities										
Public park										
Yes	7.3	6.6	8.5	6.3	4.6	11.1	10.2	7.9	6.5	5.0
Indoor recreational facilities										
Yes	0.8	0.9	0.8	0.6	0.8	1.0	1.8	0.9	0.3	0.4
Outdoor recreational facilities										
Yes	2.8	1.0	1.7	0.8	2.5	7.5	4.6	3.1	1.8	2.1

*SES, socioeconomic status.

t-distribution, values greater than |1.96| correspond to a *p*-value less than 0.05. Then, associations between clusters and regions were assessed with Fisher's exact tests and the relations between categorical variables (e.g. between one cluster and one specific region) were examined. Tests are based on the comparison of the proportion of individuals who possess the second category (e.g. one region) among those who possess the first (e.g. the first cluster) and the global percentage of individuals who possess the second category (the first cluster). Strength of associations is given by *v*-test values, following a Student's *t*-distribution.

All analyses were performed using the FactoMineR package (32) from R software, version 3.1.1 (33).

Results

Overall description of Google Street View-virtual audit data

Table 1 describes the virtual audit data. Percentages of street segments for each of the 56 built environmental features are shown by European regions and by type of neighbourhoods. General observations included are that (i) some features were found only in very few street segments (<1%). This was the case for instance for 'segments with shared cycle path with pedestrians', 'presence of railway/underground stations', 'presence of street food markets', 'presence of wine/liquor stores', 'presence of on-street food vendors', 'presence of shopping malls' or 'presence of indoor recreational facilities'. (ii) There were some important between-country differences. For instance, sidewalk-related features (in the domain 'walking') varied substantially from one country to the other (e.g. 14.6% of Hungarian segments included well-maintained sidewalks, compared with 68.9% in the Netherlands).

Results from multiple factor analysis

The first six dimensions of the MFA had eigenvalues higher than 1, but a jump in the scree plot between the third and the fourth dimension meant that only the first three dimensions were retained. These accounted for 39% of total variance (22.1%, 9.0% and 7.9%, respectively). The first dimension was mostly correlated with food environment (contribution of items related to grocery stores and food outlets to the inertia of the dimension = 31%), land use items (18.9%) and to a less extent with aesthetics (14.2%) and walkability (11%)/cyclability (11.8%) items. For example, variables most positively correlated with this dimension were the percentages of streets with cafés, restaurants, grocery stores, supermarkets, non-residential buildings and bicycle lanes, while those most negatively correlated were percentages of streets with residential gardens and trees (Supporting Information Table 1). Thus, a high value for this first dimension indicates a high level of food environment facilities and land use mix, as well as a high level of walkability and cyclability.

The second dimension was strongly associated with recreational facilities (30.1%), especially outdoor recreational facilities, and then to cyclability (25.7%) and walkability (18.7%) (Supporting Information Table 1). This can be considered as an indicator of the recreational facility contribution to the physical activity environment. The third dimension was correlated with walkability (33.1%), cyclability (24.6%) and aesthetics items (21%). Therefore, this indicates support for active travel.

Results from hierarchical clustering

Hierarchical clustering based on these first three dimensions identified four clusters of neighbourhoods (Figs 1 and 2). Overall, clusters 1 and 2 mainly encompassed low residential density neighbourhoods (Fig. 2), while clusters 3 and 4 exclusively included neighbourhoods with high residential density. There was no association between neighbourhood clusters and SES levels. Cluster 1 included more than half of neighbourhoods ($n = 33/59$). This cluster was significantly associated with variables such as percentage of streets with residential gardens, trees or green areas (Supporting Information Table 2). Moreover, this cluster was significantly associated with the category 'low residential density' (v-test = 2.44, $p < 0.05$), regardless of the neighbourhood SES level. This cluster thus defined 'green neighbourhoods with low residential density'.

Cluster 2 included 16 neighbourhoods mainly with low residential density and was associated with the third dimension, consisting of active mobility-related items, such as traffic calming devices, zebra crossings, well-maintained sidewalks and traffic lights (Supporting Information Table 2). This was confirmed by the significant association found between this cluster and the third MFA dimension (v-test = 2.41, $p < 0.05$). This cluster therefore defined 'neighbourhoods supportive of active mobility'.

Clusters 3 and 4 included only high residential density neighbourhoods. Cluster 3 included 7 neighbourhoods and was associated with food and recreational facility variables and with public bicycle and public transport facilities, confirming that it captured 'high residential density neighbourhoods with food and recreational facilities'.

Cluster 4 included only three neighbourhoods because these differed strongly from all other neighbourhoods. It was related to food environment and walkability/cyclability items, but also to graffiti and abandoned buildings (Supporting Information Table 2). This cluster thus defined 'high residential density neighbourhoods with low level of aesthetics'.

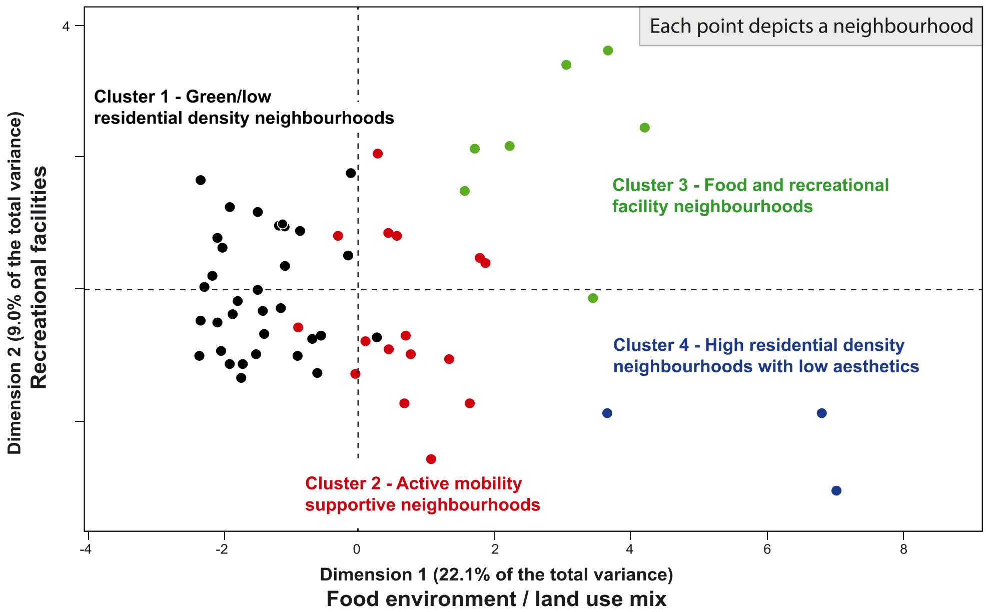

Figure 1 Distribution of the 59 neighbourhoods audited in the SPOTLIGHT project in the factorial map and identification of the derived clusters. Clusters were identified by hierarchical clustering on principal components (i.e. the first three dimensions of multiple factor analysis).

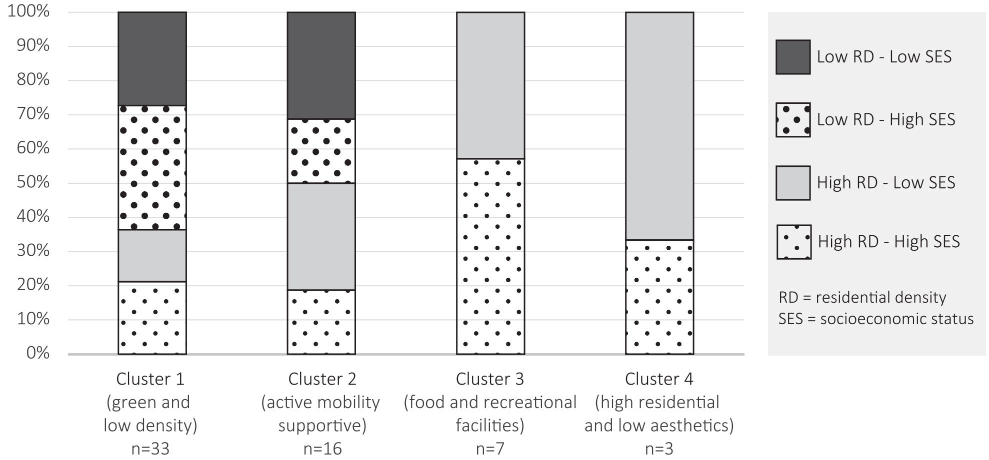

Figure 2 Frequency of SPOTLIGHT neighbourhood types by identified clusters.

Relations between clusters, types of neighbourhoods and European regions

Overall, distribution of the four clusters varied significantly across European regions (Fisher's exact test $p < 0.001$). Cluster 1 was underrepresented in Paris and suburbs (v-test = −2.96, $p < 0.01$), cluster 2 was overrepresented in Paris and suburbs (v-test = 2.48, $p < 0.05$) but underrepresented in Greater London (v-test = −2.46, $p < 0.05$), cluster 3 was overrepresented in Greater London (v-test = 3.82, $p < 0.001$) and cluster 4 overrepresented in Paris area, because the three neighbourhoods on this cluster were all in this region.

The distribution and scope of clusters by country (Fig. 3) reveals three cluster types in the Budapest and Paris areas, whereas only two types emerged in Ghent, the Randstad and London areas.

Finally, we noted that clusters 3 and 4 only include high residential density neighbourhoods, while both levels of neighbourhood density are represented in clusters 1 and 2.

Discussion

In this study, we identified four clusters of neighbourhoods using a comprehensive environmental dataset obtained

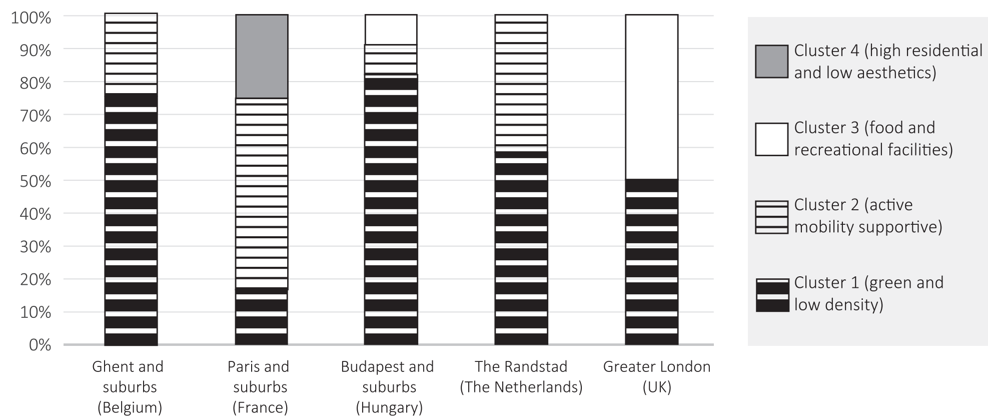

Figure 3 Frequency of the four identified clusters by each urban region studied in the SPOTLIGHT project.

through remote sensing by Google Street View in 59 neighbourhoods of five European urban zones. Using cluster analysis, a neighbourhood typology was identified based on (often co-existing) food and physical activity-related neighbourhood features. Moreover, given this was a rich dataset with hierarchical structure (items grouped into domains), MFA was used to balance the influence of groups of variables and therefore to refine the clustering of neighbourhoods.

The first main finding is the identification of four clusters of neighbourhoods differing in obesogenic characteristics and with heterogeneous structures. A key point is that, in some clusters, features co-exist that are potentially obesogenic as well as non-obesogenic. For example, presence of cycle lanes and public bicycle facilities is expected to influence active mobility – and thus overall physical activity – but may share space with impediments to active mobility (e.g. low levels of aesthetics). This typically concerns high residential density neighbourhoods with high level of facilities (clusters 3 and 4). Greener neighbourhoods with high levels of aesthetics and low residential density – hence potentially fostering recreational walking and cycling – are also those with a low presence of active transport facilities (e.g. no sidewalk, no bike lane, etc.) and lower density of destinations, and may therefore be more oriented towards motorized transport. Along the same rationale, high residential density neighbourhoods that provide infrastructure to support active transport (clusters 3 and 4) are also those where obesogenic food outlets (e.g. fast food restaurants) are the most available. These findings suggest that different obesogenic dimensions of the built environment should be considered together, as counteracting features may exist in the same neighbourhood (34). We should therefore recognize the multifaceted spatial interactions between environmental features rather than seeing them in isolation (35). This will capture place-specific dynamics that structure the potential relation between built environment, obesity-related behaviours and obesity, often oversimplified in existing literature (9). This corroborates the conclusions of previous studies addressing obesogenic environments in urban contexts (4,34) and shows the importance of factor and cluster analyses to quantify such multifactorial, complex local contexts (11–13,36–39).

The second main finding is that built environmental clusters derived from factor analysis were unequally represented and distributed across European regions. Study neighbourhoods in Greater London differed from those in the Paris area (differences between regions), in terms of environmental features, despite a sampling strategy that randomly selected neighbourhoods to be representative. Neighbourhoods in the Randstad, Ghent and Budapest areas were more similar (similitudes between regions). Neighbourhoods in the London region were mostly characterized by a high level of recreational facilities (especially public parks and outdoor recreational facilities, Table 1), while Paris neighbourhoods were characterized by high urban densities and a strong presence of food outlets (cafés, restaurants, supermarkets and local shops). The Randstad, Ghent and Budapest neighbourhoods included in the present study appeared more similar to each other, characterized by lower residential densities and greener areas, together with a very low percentage of streets with food and recreational facility items compared with neighbourhoods selected in London and Paris.

International differences in built environment characteristics are consistent with results from the International Physical Activity and the Environment Network (IPEN) study (40). IPEN showed that measures of environmental features comparable with those used in this study (although GIS-derived and specifically based on physical activity features) in 11 countries led to profound between-country differences in urban form (41). As mentioned, periods of development, topography, economic conditions, cultural norms, municipal zoning laws and public health planning practices are possible explanations for these between-country differences.

Potential limitations of this study relate to those of virtual audit-based assessment (20), namely, (i) potential sources of error because of imagery date disruption, limitations and temporal variability of neighbourhood features (18), (ii) non-complete GSV coverage, especially for small pedestrian streets (16) and (iii) potential inaccuracies when assessing small, micro-environmental items (e.g. litter or garbage), which may be difficult for an assessor to discern and which may thus require on-site ratings (17,42).

The main strength of the current study is that it is based on objective assessment of environments, following the same protocol across countries. The pre-stratification was made on the basis of key neighbourhood characteristics related to obesity-related behaviours (SES/residential density). Finally, the factor and cluster analysis techniques we used (MFA) were rather novel and particularly suited to our data structure.

Multidimensional clusters of neighbourhoods are commonly used in marketing activities (43), and we believe that they may help policymakers better address the obesity issue through urban planning interventions. This study suggests that interventions should be tailored according to the type of neighbourhoods (varying as a function of the specific combination of environmental features they reveal), rather than to the presence of specific items or to the country studied. In other words, the local context plays a central role. Previous research has already highlighted the importance of locally tailored interventions, especially for obesity-related behaviours (38,44–48), justifying the neighbourhood as the basis for intervention in many obesity-related studies (49). This study suggests the importance of taking into account the overall context rather than planning or modifying isolated environmental features in future analyses relating the built environment to obesity-related outcomes. Consequently, it also suggests that recurrent inconsistencies of associations between the environment and obesity-related behaviours noted in the literature may come from these varying types of neighbourhoods. Policymakers should take a comprehensive view of the multiple facets (and their interactions) of local environments, thinking locally within a global framework, even if it may appear more challenging. As recommended by Heath *et al.* (50), local needs assessments, involving both qualitative and quantitative investigations, should inform the design of interventions and help to ensure that they genuinely help local communities adopt more healthy lifestyles.

In conclusion, in this study, virtual audit data were used to identify four clusters, reflecting the complexity of obesogenic food and physical activity features in the built environment in different neighbourhoods. The distribution of these patterns varied across urban regions in different countries across Europe. These findings suggest that multidimensional constructs of the built environment should be targeted rather than isolated features, because these are likely to interact differently in each local context.

Declaration of interests

The authors have no conflicts of interest to declare.

Acknowledgements

The SPOTLIGHT project was funded by the Seventh Framework Programme (CORDIS FP7) of the European Commission, HEALTH (FP7-HEALTH-2011-two-stage), grant agreement no. 278186. The content of this article reflects only the authors' views, and the European Commission is not liable for any use that may be made of the information contained therein.

Supporting information

Additional Supporting Information may be found in the online version of this article, http://dx.doi.org/10.1111/obr.12378

Table S1. Pearson coefficient correlations between the initial environmental features and the first three dimensions derived from multiple factor analysis. Neighbourhood-level environmental feature variables (each row in the Table) are expressed as the percentage of street segments audited including at least one given feature.

Table S2. Associations between the initial environmental features and the four clusters (v-test values $>|1.96|$). The higher the v-test, the more the feature characterizes the cluster. Neighbourhood-level environmental feature variables (each row in the Table) are expressed as the percentage of street segments audited including at least one given feature.

References

1. Sallis JF, Owen N, Fisher EB. Ecological models of health behavior. In: Health Behavior and Health Education: Theory, Research, and Practice. Jossey-Bass: Wiley, 2008, pp. 465–482.

2. Pikora TJ, Bull FCL, Jamrozik K, Knuiman M, Giles-Corti B, Donovan RJ. Developing a reliable audit instrument to measure the physical environment for physical activity. *Am J Prev Med* 2002; **23**: 187–194.

3. Renalds A, Smith TH, Hale PJ. A systematic review of built environment and health. *Fam Community Health* 2010; **33**: 68–78.

4. Economos CD, Hatfield DP, King AC, Ayala GX, Ann PM. Food and physical activity environments: an energy balance approach for research and practice. *Am J Prev Med* 2015; **48**: 620–629.

5. Butland B, Jebb S, Kopelman P, et al. Foresight. Tackling obesities: future choices. Project report. 2007; 155 pp.

6. Swinburn B, Egger G, Raza F. Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Prev Med* 1999; **29**: 563–570.

7. Lake A, Townshend TG, Alvanides S (eds.). Obesogenic Environments: Complexities, Perceptions and Objective Measures. Wiley-Blackwell: Oxford, 2010.

8. Mackenbach JD, Rutter H, Compernolle S *et al.* Obesogenic environments: a systematic review of the association between the

physical environment and adult weight status, the SPOTLIGHT project. *BMC Public Health* 2014; **14**: 233.

9. Burgoine T, Alvanides S, Lake AA. Creating 'obesogenic realities'; do our methodological choices make a difference when measuring the food environment? *Int J Health Geogr* 2013; **12**: 33.

10. Nelson MC, Gordon-Larsen P, Song Y, Popkin BM. Built and social environments associations with adolescent overweight and activity. *Am J Prev Med* 2006; **31**: 109–117.

11. Boone-Heinonen J, Evenson KR, Song Y, Gordon-Larsen P. Built and socioeconomic environments: patterning and associations with physical activity in U.S. adolescents. *Int J Behav Nutr Phys Act* 2010; **7**: 45.

12. Charreire H, Weber C, Chaix B *et al*. Identifying built environmental patterns using cluster analysis and GIS: relationships with walking, cycling and body mass index in French adults. *Int J Behav Nutr Phys Act* 2012; **9**: 59.

13. Adams MA, Ding D, Sallis JF *et al*. Patterns of neighborhood environment attributes related to physical activity across 11 countries: a latent class analysis. *Int J Behav Nutr Phys Act* 2013; **10**: 34.

14. Zenk SN, Powell LM, Rimkus L *et al*. Relative and absolute availability of healthier food and beverage alternatives across communities in the United States. *Am J Public Health* 2014; **104**: 2170–2178.

15. Papas MA, Alberg AJ, Ewing R, Helzlsouer KJ, Gary TL, Klassen AC. The built environment and obesity. *Epidemiol Rev* 2007; **29**: 129–143.

16. Clarke P, Ailshire J, Melendez R, Bader M, Morenoff J. Using Google earth to conduct a neighborhood audit: reliability of a virtual audit instrument. *Health Place* 2010; **16**: 1224–1229.

17. Rundle AG, Bader MDM, Richards CA, Neckerman KM, Teitler JO. Using Google Street View to audit neighborhood environments. *Am J Prev Med* 2011; **40**: 94–100.

18. Curtis JW, Curtis A, Mapes J, Szell AB, Cinderich A. Using Google Street View for systematic observation of the built environment: analysis of spatio-temporal instability of imagery dates. *Int J Health Geogr* 2013; **12**: 53.

19. Kelly CM, Wilson JS, Baker EA, Miller DK, Schootman M. Using Google Street View to audit the built environment: inter-rater reliability results. *Ann Behav Med* 2012; **45**: 108–112.

20. Charreire H, Mackenbach JD, Ouasti M *et al*. Using remote sensing to define environmental characteristics related to physical activity and dietary behaviours: a systematic review (the SPOTLIGHT project). *Health Place* 2014; **25**: 1–9.

21. Silva V, Grande AJ, Rech CR, Peccin MS. Geoprocessing via Google maps for assessing obesogenic built environments related to physical activity and chronic noncommunicable diseases: validity and reliability. *J Healthc Eng* 2015; **6**: 41–54.

22. Bethlehem JR, Mackenbach JD, Ben Rebah M *et al*. The SPOTLIGHT virtual audit tool: a valid and reliable tool to assess obesogenic characteristics of the built environment. *Int J Health Geogr* 2014; **13**: 52.

23. Lakerveld J, Brug J, Bot S *et al*. Sustainable prevention of obesity through integrated strategies: the SPOTLIGHT project's conceptual framework and design. *BMC Public Health* 2012; **12**: 793.

24. Lakerveld J, Ben Rebah M, Mackenbach JD *et al*. Obesity-related behaviours and BMI in five urban regions across Europe: sampling design and results from the SPOTLIGHT cross-sectional survey. *BMJ Open* 2015; **5** e008505.

25. Demarest S, Van der Heyden J, Charafeddine R, Tafforeau J, Van Oyen H, Van Hal G. Socio-economic differences in participation of households in a Belgian national health survey. *Eur J Public Health* 2013; **23**: 981–985.

26. Escofier B, Pagès J. Analyses factorielles simples et multiples: Objectifs, méthodes et interprétation. 4e édition. Dunod: Paris, 2008.

27. Li Y-S, Chuang Y-C. Neighborhood effects on an individual's health using neighborhood measurements developed by factor analysis and cluster analysis. *J Urban Health Bull N Y Acad Med* 2009; **86**: 5–18.

28. Cervero R, Sarmiento OL, Jacoby E, Gomez LF, Neiman A. Influences of built environments on walking and cycling: lessons from Bogotá. *Int J Sustain Transp* 2009; **3**: 203–226.

29. Rodríguez DA, Brisson EM, Estupiñán N. The relationship between segment-level built environment attributes and pedestrian activity around Bogota's BRT stations. *Transp Res Part Transp Environ* 2009; **14**: 470–478.

30. Bécue-Bertaut M, Pagès J. Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. *Comput Stat Data Anal* 2008; **52**: 3255–3268.

31. Husson F, Lê S, Pagès J. Exploratory Multivariate Analysis by Example Using R. CRC Press: Boca Raton, 2011.

32. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. *J Stat Softw* 2008; **15**.

33. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2015.http://www.R-project.org/.

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

35. Burgoine T, Alvanides S, Lake AA. Assessing the obesogenic environment of North East England. *Health Place* 2011; **17**: 738–747.

36. Adams MA, Sallis JF, Kerr J *et al*. Neighborhood environment profiles related to physical activity and weight status: a latent profile analysis. *Prev Med* 2011; **52**: 326–331.

37. Wall MM, Larson NI, Forsyth A *et al*. Patterns of obesogenic neighborhood features and adolescent weight: a comparison of statistical approaches. *Am J Prev Med* 2012; **42**: e65–e75.

38. 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.

39. Kurka JM, Adams MA, Todd M *et al*. Patterns of neighborhood environment attributes in relation to children's physical activity. *Health Place* 2015; **34**: 164–170.

40. Kerr J, Sallis JF, Owen N *et al*. Advancing science and policy through a coordinated international study of physical activity and built environments: IPEN adult methods. *J Phys Act Health* 2013; **10**: 581–601.

41. Adams MA, Frank LD, Schipperijn J *et al*. International variation in neighborhood walkability, transit, and recreation environments using geographic information systems: the IPEN adult study. *Int J Health Geogr* 2014; **13**: 43.

42. Vanwolleghem G, Dyck DV, Ducheyne F, Bourdeaudhuij ID, Cardon G. Assessing the environmental characteristics of cycling routes to school: a study on the reliability and validity of a Google Street View-based audit. *Int J Health Geogr* 2014; **13**: 19.

43. Yorke DA, Meehan SA. Acorn in the political marketplace. *Eur J Mark* 1986; **20**: 63–76.

44. Lopez RP, Hynes HP. Obesity, physical activity, and the urban environment: public health research needs. *Environ Health* 2006; **5**: 25.

45. Leslie E, Coffee N, Frank L, Owen N, Bauman A, Hugo G. Walkability of local communities: using geographic information

systems to objectively assess relevant environmental attributes. *Health Place* 2007; **13**: 111–122.
46. Stafford M, Cummins S, Ellaway A, Sacker A, Wiggins RD, Macintyre S. Pathways to obesity: Identifying local, modifiable determinants of physical activity and diet. *Soc Sci Med* 2007; **65**: 1882–1897.
47. Santana P, Santos R, Nogueira H. The link between local environment and obesity: a multilevel analysis in the Lisbon metropolitan area, Portugal. *Soc Sci Med* 2009; **68**: 601–609.
48. Huang R, Moudon AV, Cook AJ, Drewnowski A. The spatial clustering of obesity: does the built environment matter? *J Hum Nutr Diet* 2014; **28**(6): 604–612.
49. Black JL, Macinko J. Neighborhoods and obesity. *Nutr Rev* 2008; **66**: 2–20.
50. Heath GW, Brownson RC, Kruger J, Miles R, Powell K, Ramsey L. The effectiveness of urban design and land use and transport policies and practicies to increase physical activity: a systematic review. *J Phys Act Health* 2006; **3**: S55–S76.