LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Charreire, H; Feuillet, T; Roda, C; Mackenbach, JD; Compernolle, S; Glonti, K; Brdos, H; Le Vaillant, M; Rutter, H; McKee, M; De Bourdeaudhuij, I; Brug, J; Lakerveld, J; Oppert, JM (2015) Self-defined residential neighbourhoods: size variations and correlates across five European urban regions. Obesity reviews, 17 Suppl 1. pp. 9-18. ISSN 1467-7881 DOI: https://doi.org/10.1111/obr.12380

DOI: 10.1111/obr.12380

Supplement Article

Self-defined residential neighbourhoods: size variations and correlates across five European urban regions

H. Charreire,[1,2] T. Feuillet,[1] C. Roda,[1] J. D. Mackenbach,[3] S. Compernolle,[4] K. Glonti,[5] H. Bárdos,[6] M. Le Vaillant,[7] H. Rutter,[5] M. McKee,[5] I. De Bourdeaudhuij,[4] J. Brug,[3] J. Lakerveld[3] and J.-M. Oppert[1,8]

[1]Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Université Paris 13, Centre de Recherche en Epidémiologie et Statistiques, Inserm (U1153), Inra (U1125), Cnam, COMUE Sorbonne Paris Cité, 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 Centre, 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, Hungary; [7]CERMES3, UMR 8211-U988, CNRS, INSERM, Université Paris Descartes, EHESS, Villejuif, France; [8]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: H. Charreire, Paris Est University, UPEC, Lab-Urba, Urban School of Paris, Créteil 94010, France.
E-mail: helene.charreire@u-pec.fr

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

Summary

The neighbourhood is recognized as an important unit of analysis in research on the relation between obesogenic environments and development of obesity. One important challenge is to define the limits of the residential neighbourhood, as perceived by study participants themselves, in order to improve our understanding of the interaction between contextual features and patterns of obesity. An innovative tool was developed in the framework of the SPOTLIGHT project to identify the boundaries of neighbourhoods as defined by participants in five European urban regions. The aims of this study were (i) to describe self-defined neighbourhood (size and overlap with predefined residential area) according to the characteristics of the sampling administrative neighbourhoods (residential density and socioeconomic status) within the five study regions and (ii) to determine which individual or/and environmental factors are associated with variations in size of self-defined neighbourhoods. Self-defined neighbourhood size varies according to both individual factors (age, educational level, length of residence and attachment to neighbourhood) and contextual factors. These findings have consequences for how residential neighbourhoods are defined and operationalized and can inform how self-defined neighbourhoods may be used in research on associations between contextual characteristics and health outcomes such as obesity.

Keywords: Multilevel, overlap, self-defined neighbourhoods, SPOTLIGHT.

Abbreviations: BMI, body mass index; GIS, geographical information systems; SES; socio economic status.

obesity reviews (2016) **17** (Suppl. 1), 9–18

Introduction

Obesity is the outcome of complex interactions between a large number of factors related both to the individual-level and to contextual-level dimensions (1--4). In this field of research, neighbourhood is increasingly recognized as one important unit of analysis in research on health behaviors (5) and obesity (6,7). There is no agreed definition of the term 'neighbourhood' in population health and urban planning research (8--10). Two broad approaches have been used. The first is to select predefined areas, such as those defined by administrative boundaries (e.g. census tract) or buffer zones centred on home addresses, although with differing radii. The second approach recognizes that predefined areas do not necessarily represent what residents perceive as their neighbourhood (11,12), but this leaves the challenge of how to define and measure self-defined neighbourhoods (12--16). A small but growing literature has proposed methods to delineate self-defined neighbourhood limits in health-related studies (13--17). However, so far, most studies consist of small samples, with fewer than 70 participants (13,14,16--18). Traditional paper maps on which subjects were asked to draw the limits of the area they perceive as their neighbourhood were commonly used (11--14,17,18). Only one pilot study has used a web-based tool (16), a potentially cost-effective way of collecting data in a large population study. The geographical generalizability of this literature is uncertain; we identified only two studies performed in Europe (13,15), whereas most other studies were performed in the USA (11,12,14,17--21). Moreover, no study compared findings in different countries, so we lack information on how the congruence, or overlap, between self-defined neighbourhoods and predefined areas varies. Finally, self-definition of neighbourhoods has been found to be related to individual characteristics including age, gender, educational level or length of residence in the neighbourhood (12,19,20). However, few studies have investigated whether contextual characteristics such as population density or socioeconomic status (SES) also influence how the limits of neighbourhoods are defined (12,15,19).

Within the framework of the EU-funded Spotlight project, which aims to provide a comprehensive overview of multiple levels of determinants of overweight and obesity in order to inform effective obesity prevention approaches (22), a web-based tool was developed to enable participants in five European urban regions to self-define their neighbourhood limits. In this paper, we (i) describe the characteristics of self-defined neighbourhoods (i.e. size and overlap with predefined residential area) in relation to the characteristics of the sampling administrative neighbourhoods (residential density and SES) within the five study regions and (ii) explored which individual (age, gender, educational level, employment status, length of residential neighbourhood and BMI) or/and contextual determinants (neighbourhood residential density and SES and study regions) are associated with the variation in size of self-defined neighbourhoods.

Methods

Study design and sampling

A total of 60 administrative neighbourhoods were selected in five urban regions across Europe (i.e. 12 administrative neighbourhoods per region): Ghent and suburbs (Belgium), Paris and inner suburbs (France), Budapest and suburbs (Hungary), the Randstad (including the cities of Amsterdam, Rotterdam, The Hague and Utrecht in the Netherlands) and Greater London (UK). As described in detail elsewhere (23), sampling of administrative neighbourhoods was based on two environmental obesity-related characteristics to ensure variability of environmental contexts: residential density and neighbourhood median income. Residential density data were obtained from the Urban Atlas database distributed by the European Environmental Agency, based on a compilation of satellite photographs covering Europe providing high-resolution data on land use data (24). This Urban Atlas includes a measure of density of residential areas (calculated as the percentage of coverage of buildings devoted to residential facilities) that is comparable across European countries. Data on residential density on the SPOTLIGHT project were defined using two classes – high residential and low residential density – corresponding to >80% and <50% of areas covered by residential buildings, respectively. Median income data were derived from national census databases from all five countries, with two classes used: low and high (i.e. the first and third tertiles). The combination of residential density and neighbourhood-level income classes allowed four neighbourhood types to be defined (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. We anticipated that response rates would vary according to neighbourhood SES so sampled 1,200 adults in low SES neighbourhoods and 800 in high SES neighbourhoods. Neighbourhoods were randomly selected from within three different administrative areas in each country, generating a total of 60 neighbourhoods.

Study participants and questionnaire

Adult inhabitants (≥18 years) within the 60 selected administrative neighbourhoods were invited to participate in an online survey with a similar approach in each country. The study design and sampling approach have been described previously (23). Briefly, in the selected neighbourhoods, a random sample of residential addresses was drawn from postal companies (the Netherlands), Yellow Pages (France), electoral rolls (UK) or public administration services (Belgium and Hungary). Between February and September 2014, participants were recruited via postal invitation using the Dillman method (25). A total of 6,037 (10.8%, out of 55,893) individuals participated in the study. The online survey included questions on demographics and neighbourhood perceptions, as well as weight and height. In addition, a web-mapping tool allowing respondents to draw the limits of their self-defined neighbourhood was developed and included in the survey questionnaire. In the present study, we used data from participants who provided data on self-defined neighbourhood boundaries ($N = 4{,}454$). The study was approved by the corresponding local ethics committees of participating countries, and all participants in the survey provided informed consent.

Self-defined neighbourhood limits and overlap

Participants were asked to draw what they considered as their neighbourhood limits on an online open layer map centred on their residential address, following the instruction '*Please draw the boundaries of what you consider as your neighbourhood on the map below*'. Using their computer mouse or trackpad, subjects clicked to create points on the map at the borders of their perceived neighbourhoods. All neighbourhood geographical coordinate points (longitude and latitude) were recorded as feature attributes in a geographic information system (GIS). All the points were combined to form an enclosed area (polygon boundaries) representing the self-defined neighbourhood of a given participant. The size of each self-defined neighbourhood was calculated in square kilometres. Self-defined neighbourhoods with a size over 40 km^2 (provided by 80 participants) were excluded. Analyses in the present study were thus based on data from 4,374 participants in the five European urban regions.

The geographic information system was also used to geolocalize home addresses of participants and to define a Euclidean buffer with a radius of 500 m centred on each participant's home. Although the optimal buffer size in this type of research has not been clearly established (26), a similar buffer (i.e. 500 m corresponding, on average, to 6 to 10-min walk) has been used in a recent international study (IPEN Adult study (26)). This step was performed only for participants with residential addresses that could be geolocalized (3,621 participants). The proportion (%) of the self-defined neighbourhood area covered by the buffer area was computed to explore any overlap in area size (Fig. 1).

Individual correlates of self-defined neighbourhood size

Individual characteristics of participants included gender, age (input in models as a continuous variable and centred on the mean), educational level (input as a dichotomous variable defined in each country by lower [from less than primary to higher secondary education] and higher level

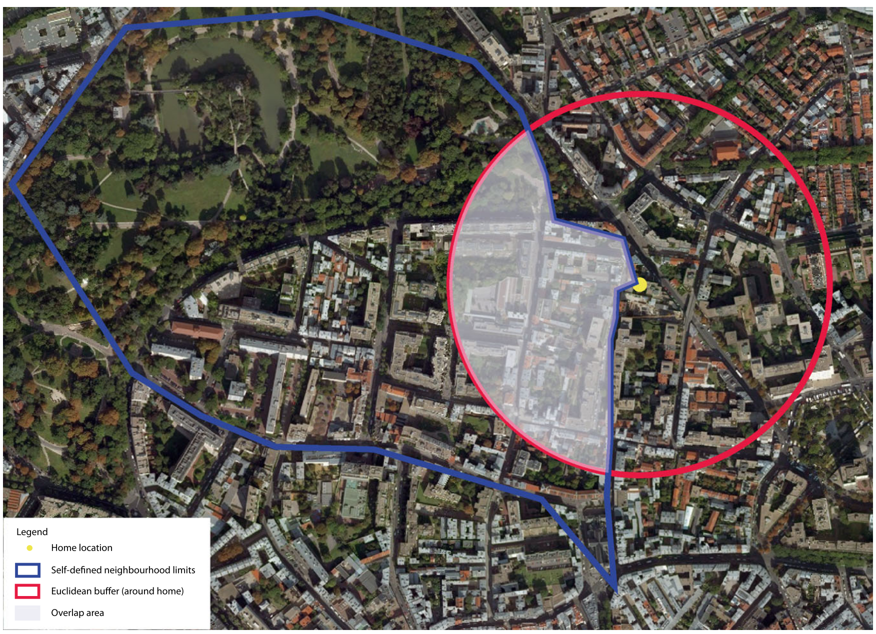

Figure 1 Overlap area between self-defined neighbourhood and buffer around home – an example in Paris (France).

[college or university level] allowing comparison between country-specific education systems), employment status (currently employed, currently not employed, retired, in education and homemaker) and BMI (as a continuous variable, centred on the mean, calculated by dividing self-reported weight in kilogrammes by the square of the self-reported height in metres). Participants were asked about how attached they felt to their neighbourhood by selecting which of the following statements best reflected their current residential neighbourhood situation: I would like to (i) continue to live in my neighbourhood, (ii) move to another neighbourhood or (iii) no specific wish about moving. Respondents were also asked about the number of years they had lived in their neighbourhood (input as a dichotomous variable defined by the median value: less than 10 years and 10 years or more).

Self-defined neighbourhood analysis

In order to analyse self-defined neighbourhood differences in size and overlap across neighbourhoods and study regions, descriptive statistics such as percentages, maximum, median and mean with standard deviation (SD) were used to summarize participant characteristics, self-defined neighbourhood size and percent overlap distributions.

In order to explore determinants (at both individual and contextual levels) of size variation and according to both the distribution of the dependant variable and the hierarchical structure of the data (individuals nested within administrative neighbourhoods nested within countries), we applied three-level Poisson regression analysis, with the levels being the individual, the administrative residential neighbourhood and the region. Given the non-Gaussian, zero-inflated distribution of the dependent variable (size of the self-defined neighbourhoods), each value was rounded to the nearest half-unit (0.5). This discretization procedure enabled the variable to be modelled using a Poisson regression. Models were implemented using individual-level characteristics of participants (age, gender, educational level, employment status, length of residence, BMI and neighbourhood attachment) at level 1. Contextual characteristics (residential density and SES levels of administrative neighbourhood of residence) were included at level 2 and the region of residence at level 3. First, an 'empty' or null model (i.e. one that only includes a random intercept) was created to assess the components of variance within and between administrative neighbourhoods and European regions. The individual characteristics were then added (model 1). Finally, the characteristics of administrative neighbourhoods (residential density and SES) were included (model 2). The variation of the component variances between models was examined. The amount of the variance explained (between the null model and both models 1 and 2) was calculated by the proportional change in variance (PCV). Interaction between residential density and SES levels on size of self-defined neighbourhoods was examined but was not included in the model because of non-significance ($p = 0.07$).

There were approximately 13% cases with missing data on at least one study covariate. Assuming that data were missing at random (i.e. the probability that a variable is missing was not related to other observed data), 15 imputed datasets were created, as recommended by Rubin (27) and Bodner (28). Multiple imputations were performed using chained equations (29). Continuous variables were imputed using predictive mean matching with their five nearest neighbours, and categorical variables were imputed using logistic regression. Variables entered in imputation models were those required for the planned analyses. All analyses were performed with SAS version 9.3 (SAS Institute Inc., Cary, NC, USA)

Results

Table 1 presents the characteristics of the study sample. The mean (SD) age was 50.8 (16.2) years, and 55.3% of subjects were women. Most participants had lived in their residential neighbourhood for 10 years or more (62.7%) or indicated that they '*would like to stay in this neighbourhood*' (66.2%). Except for gender, socio-demographic differences between participants who reported self-defined neighbourhood and participants who did not report self-defined neighbourhoods were observed (Table 1). Subjects who did not report self-defined neighbourhood limits were older and had a higher BMI, and a higher proportion of them had lived for 10 years or more in the neighbourhood, had a low educational level, were currently not employed and were from low residential density/low SES neighbourhoods.

Descriptions of size and overlap across neighbourhoods and study regions are presented in Table 2. On average, the size (SD) of self-defined neighbourhood was 1.96 (4.00) km^2 and differed significantly across neighbourhood types and European regions. The largest mean size was drawn by respondents from low density/high SES neighbourhoods (2.42 [4.53] km^2) and smallest mean size in high density/low SES (1.65 [3.53] km^2) neighbourhoods. The mean size was lowest in France (0.78 [1.66] km^2) and highest in Hungary (3.31 [5.53] km^2). The mean (std) proportion of self-defined neighbourhoods, which overlapped with the area of the 500 m Euclidean buffer, was 29.7 (17.9)%. The highest percent overlap was found in high density/low SES neighbourhoods and the lowest in low density/low SES neighbourhoods. At the European region level, residents from the neighbourhoods of Greater London had the highest percent overlap between self-defined neighbourhood and the buffer. The lowest percentage was observed in Ghent and suburbs.

Table 1 Characteristics of the study sample

	Participant who reported self-defined neighbourhood		Participant who did not report self-defined neighbourhood		Difference between subsamples
	N 4,374	% Mean (std)	*N* 1,663	% Mean (std)	*p**
Gender					
Male	1,940	44.7	691	42.3	0.0984
Female	2,403	55.3	943	57.7	
Missing	31		29		
Age (in years)		50.8 (16.2)		54.7 (16.8)	<0.0001
Missing	30		32		
How many years have you lived in this neighbourhood?					
Less than 10 years	1,595	37.3	509	33.5	0.0079
10 years or more	2,684	62.7	1,012	66.5	
Missing	95		142		
Educational level					
Low	1,728	42.4	802	58.9	<0.0001
High	2,351	57.6	559	41.1	
Missing	295		302		
Employment status					
Currently employed	2,562	58.7	731	44.3	<0.0001
Currently not employed	289	6.6	152	9.2	
Retired	1,135	26.0	616	37.4	
In education	172	3.9	52	3.2	
Homemaker	207	4.7	98	5.0	
Missing	9		14		
BMI (kg/m^2)		25.0 (4.3)		25.8 (4.9)	<0.0001
Missing	385		357		
Which of the following statements best reflects your current situation?					
I would like to stay in this neighbourhood	2,873	66.2	1,027	67.4	0.0259
I would like to move house	855	19.7	308	20.2	
I don't have a specific preference to move house	612	14.1	189	12.4	
Missing	34		139		
Neighbourhood patterns					
High residential density–high SES	1,107	25.3	324	20.7	0.0001
High residential density–low SES	1,021	23.3	374	23.9	
Low residential density–high SES	1,104	25.2	379	24.3	
Low residential density–low SES	1,142	26.1	485	31.1	
Regions					
Paris and suburbs (France)	689	15.8	155	9.3	<0.0001
The Randstad (the Netherlands)	1,250	28.6	359	21.6	
Ghent and suburbs (Belgium)	1,210	27.7	639	38.4	
Greater London (UK)	580	13.3	280	16.8	
Budapest and suburbs (Hungary)	645	14.8	230	13.8	

*Chi-squared test and Mann–Whitney test
BMI, body mass index; SES, socioeconomic status; std, standard deviation.

The results from the multilevel analyses to explore determinants (at both individual and contextual levels) of size variation (as dependent variable) of self-defined neighbourhoods are presented in Table 3. Among individual-level variables, age, gender, educational level, length of residence and neighbourhood attachment were significantly associated with variation in mean size. Men described a larger mean size than women, and the size of the self-defined neighbourhood decreased significantly with age. High educational level and having resided in the neighbourhood for more than 10 years were positively associated with neighbourhood size. Those who reported '*I would like to stay* (in my neighbourhood)' also described larger neighbourhoods than those who '*would like to move*' from their neighbourhood. There was no significant difference according to employment status and BMI.

With regard to contextual dimensions, a high level of residential density was significantly associated with a lower size of self-defined neighbourhood in comparison with a low level of residential density. There was no significant difference according to SES level (Table 3). Values of variance components of the three-level models showed significant differences in size between administrative neighbourhoods

Table 2 Size variation (in km²) and percent of overlap between self-defined neighbourhood and buffer (500-m radius around home address) according to European regions and neighbourhood types

		Size (km²)					% of overlap				
		n subjects	Mean	Std	Median	Max	*n* subject*	Mean	Std	Median	Max
Regions	Paris and suburbs (France)	689	0.78	1.66	0.39	23.03	558	30.8	17.8	29.8	79.3
	The Randstad (the Netherlands)	1,250	1.47	3.32	0.59	33.10	1,139	29.6	17.2	28.4	77.9
	Ghent and suburbs (Belgium)	1,210	2.21	4.27	0.77	38.97	1,094	28.1	18.0	26.3	78.8
	Greater London (UK)	580	2.38	4.17	1.14	37.35	320	33.7	17.8	33.4	79.9
	Budapest and suburbs (Hungary)	645	3.31	5.53	1.10	34.70	504	29.3	18.9	27.0	92.3
Type of neighbourhoods	High density–high SES	1,108	1.68	3.51	0.62	34.21	949	30.2	18.3	28.6	92.3
	High density–low SES	1,021	1.65	3.54	0.66	38.97	846	32.4	18.6	32.9	81.3
	Low density–high SES	1,104	2.42	4.54	0.98	34.70	905	29.2	17.4	27.4	79.3
	Low density–low SES	1,142	2.05	4.25	0.63	37.35	915	27.2	16.9	25.6	79.9
	Total	4,374	1.96	4.00	0.71	38.97	3,621	29.7	17.9	28.4	92.3

*Excluded subject with missing data for residential address (so without buffer around residential address)
Max, maximum; *N*, number of subjects; SES, socioeconomic status; std, standard deviation.

(0.075 [0.019] in null model) but not between study regions (0.181 [0.134] in null model).

The inclusion of individual and contextual characteristics reduced the between-neighbourhood variance (variance [SE] ranged from 0.075 [0.019] in null model to 0.055 [0.016] in model 2). These results mean that 27.6% of the initial variance in the size of self-defined neighbourhoods between residential administrative neighbourhoods was explained by individual and contextual characteristics included in the model. The individual-level variables explained 7.9% of the variance (model 1), whereas contextual-level variables (without control for individual-level risk factors, data not shown) explained 19.7% of the variance. All these findings were robust to imputation of data except for the variance components between administrative neighbourhoods level, which increase (in contrast to results observed in the no imputed dataset (Supporting information Table S1)).

Discussion

The aim of this study was to investigate individual and contextual factors associated with the size of self-defined neighbourhoods in a large-scale study in five different European urban regions. The findings indicate that both individual (age, educational level, length of residence and neighbourhood attachment) and contextual factors were significantly associated with the size of self-defined neighbourhoods. The size of the self-defined neighbourhood decreased with age. Educational level and length of residence were positively associated with the size of self-defined neighbourhood. Residents who expressed a wish to stay in their neighbourhood defined larger neighbourhoods than residents who wished to move. Contextual-level factors explained substantially more variance in self-defined neighbourhood size than individual-level factors.

The use of a web-based tool in a large population in five European regions was innovative. Participants drew their neighbourhood limits following a standardized approach (23). Previous research into self-defined neighbourhoods has mainly been pilot studies in specific populations (e.g. adolescents, women and low-income communities) or cities.

The mean area of 1.96 km² (ranging from 0.80 km² in France to 3.30 km² in Hungary) in the current study is close to what was observed in pilot studies in adolescents: 3.54 km² and 1.8 km² in the study by Stewart *et al.* in Auckland (New Zealand) (16) and Robinson *et al.* in the greater Boston area (Massachusetts, USA) (17), respectively, but higher than found in previous studies in adults, where mean area sizes of less than 0.1 to 0.8 km² were reported, apart from a study conducted in 10 US cities by Coulton *et al.* (12) where an area of 2 km² was reported. In previous studies, respondents were provided with a map and asked to draw their neighbourhood during an interview with a trained technician or researcher. Numbers and types of subjects varied across studies – e.g. from 28 adolescents in Auckland, New Zealand (16) to 6,224 adults in low-income communities in 10 US cities (12). A list of response categories defining the neighbourhood of residence (i.e. ranging from 'the block street of where you live' to 'an area larger than a 15-minute walk from your house') was used among 1,630 adults in Seattle (Washington, USA) (21) and in two US studies from over 2,400 adults of the Los Angeles Family and Neighbourhood Survey (19,20). In a recent study in the Paris region, yet another method was used: 653 adults reported the name of streets or places delineating their neighbourhood and polygonal areas were built using a GIS algorithm (15).

In previous studies, no information was provided about potential exclusion criteria. Based on the principle that there are no 'right or wrong' answers to the question 'draw your neighbourhood' (16,17), we decided to define a very large area size (>40 km²) of self-defined neighbourhood as a

Table 3 Factors associated with size variation (in km²) of self-defined neighbourhood identified by multilevel Poisson regression models

	Size (km²)					
	Null model (*N* = 3,796)‡		Model 1 (*N* = 3,796)‡		Model 2 (*N* = 3,796)‡	
	Coef	(SE)	Coef	(SE)	Coef	(SE)
Fixed effects						
Individual characteristics (level 1)						
Age (in years)	–	–	−0.009***	(0.002)	−0.009***	−0.002
Gender						
Women	–	–	Ref.		Ref.	
Men	–	–	0.08**	−0.04	0.09**	−0.03
Educational level						
Low	–	–	−0.11**	(0.04)	−0.11**	(0.04)
High	–	–	Ref.		Ref.	
Employment status						
Currently employed	–	–	Ref.		Ref.	
Currently not employed	–	–	0.09	(0.07)	0.10	(0.07)
Retired	–	–	0.06	(0.06)	0.06	(0.06)
In education	–	–	0.09	(0.09)	0.09	(0.09)
Homemaker	–	–	−0.12	(0.09)	−0.12	(0.09)
BMI (kg/m^2)			−0.003	(0.004)	−0.003	(0.004)
Length of residence						
Less than 10 years	–	–	−0.19***	(0.04)	−0.18***	(0.04)
10 years or more	–	–	Ref.		Ref.	
Neighbourhood attachment						
Wish to stay in this neighbourhood	–	–	0.14**	(0.05)	0.13*	(0.05)
Wish to move	–	–	Ref.		Ref.	
No specific preference to move	–	–	0.08	(0.06)	0.07	(0.06)
Administrative neighbourhood characteristics (level 2)						
High SES	–	–	–	–	0.12	(0.07)
Low SES	–	–	–	–	Ref.	
High residential density	–	–	–	–	−0.21**	(0.07)
Low residential density	–	–	–	–	Ref.	
Between administrative neighbourhoods	0.076 (0.019)		0.070 (0.018)		0.055 (0.016)	
PCV†	–		7.90%		27.60%	
Between regions	0.181 (0.134)		0.176 (0.131)		0.175 (0.130)	
PCV†	–		5.68%		5.71%	

†[initial variance–actual variance]/initial variance: percentage of initial variance in self-defined neighbourhoods explain by the model at administrative neighbourhood level and European region level.
‡ Excluded subject with missing individual data.
*$p < 0.05$.
**$p < 0.01$.
***$p < 0.0001$. BMI, body mass index; Coef, estimated regression coefficient; PCV, proportional change in variance; SE, standard error.

criterion of exclusion based on visual outlier cut-off in all study regions. Differences in study design and data collection methodology could explain, at least in part, the observed variation in size between studies. It should be noted that our study is one of the largest of its kind, using a web-based tool that allowed data collection from a large number of subjects in different countries

Concerning individual-levels correlates of self-defined neighbourhood size, in this study, older adults defined smaller neighbourhoods than younger adults, which is consistent with previous results (12,19--21). A possible explanation could be that with increasing age, day-to-day mobility and social participation decrease, thereby reducing both self-defined neighbourhood and space within which activity is performed (30--32). Our findings also suggest that respondents with higher levels of education reported larger self-defined neighbourhoods, in line with previous research (12,19,20). One possibility is that better educated residents may experience greater mobility in terms of frequency and distance, and at the same time live in places with greater access to urban opportunities such as services, transportation and social activities (33,34). As reported in previous studies (12,19), no relation was observed between size of self-defined neighbourhood and employment status.

In our study, we also observed that, generally, women drew smaller neighbourhoods than men and that the length of residence was positively associated with the size of self-defined neighbourhood. The association of gender and length of residence with self-defined neighbourhood area size in previous studies is inconsistent. While gender differences were observed in one previous study (21), no significant association was found in others (12,18,20). As in our study, larger self-defined neighbourhoods were also reported among long-term residents in multilevel analyses of data from US adults (12,19). This result could be partially explained by the fact that longer residential length was associated with more social activities and relationships in the residential neighbourhood (35), or higher awareness of destinations, which may take a while to discover. However, other studies did not report significant associations (18) or negative associations (20,21) between length of residence and size of self-defined neighbourhood.

Although the explained variance in size of self-defined neighbourhood remained relatively low (26.7%) in our model, one of the new findings in our study was that a proportion of variance in size of self-defined neighbourhoods was explained by characteristics of the residential administrative neighbourhood and especially by residential density level. A larger self-defined neighbourhood was observed in areas characterized by low residential density compared with high residential density. Previous findings in an urban context in the USA showed smaller perceived neighbourhoods in high-density inner-city areas among adult residents in 10 cities (12) and in Green Bay, Wisconsin (18). Respondents in high-density areas are likely to reside within short distances of services, limiting the time to travel. However, other studies have found opposite results, including a French study (15) where residents of peripheral (defined by municipality population size) neighbourhoods had smaller perceived neighbourhoods than residents in inner-city (i.e. Paris municipality) neighbourhoods. In contrast, a non-significant relation between size and population density was found in US studies (19,20).

In addition, the three-level model provided evidence that the effect of the study region as included in the model was limited. However, further work is required to confirm this finding and to investigate specific contextual characteristics in each European region and their interaction with individual characteristics that may influence the definition of neighbourhood. In line with previous findings (12), additional characteristics such as variables at household level (e.g. number of children) and/or at individual level (e.g. social capital) could be added to the model. According to Kearns and Parkinson (36), neighbourhood is a complex and a multilayered phenomenon '*affected by the physical and social composition of localities – i.e. it is culturally and regionally specific*', and therefore, '*the significance of neighbourhood for different social groups varies between nations and regions*'.

The hypothesis that there would be a risk of bias arising from measures of exposure based on how residential neighbourhoods are defined has been assessed by measuring the degree of spatial congruence between self-defined neighbourhood and predefined neighbourhood (i.e. based on buffer limits) for each resident. In order to compare results across countries, a predefined neigbhourhood was delimited by a Euclidean buffer with a radius of 500 m (i.e. area of 0.79 km^2) corresponding, on average, to a 6 to 10-min walk and close to the median size of self-defined neighbourhoods in our study sample (median = 0.71 km^2). In our study, the mean percentage of overlap was around 30%, suggesting only limited spatial congruence. In agreement with these results, mean percentage (std) of perceived neighbourhood area represented 16.0 (20.0)% and 34.6 (21.7)% of the 1 mi (i.e. 1.6 km^2) Euclidean buffer area in studies in England (based on 58 adults, (13)) and New Zealand (based on 28 adolescents, (16)), respectively. In a study of 31 adolescents in greater Boston (Massachusetts, USA), no significant difference was observed between the mean size of self-defined neighbourhood (1.8 km^2) and census-defined neighbourhood (1.3 km^2). However, the authors noted that only 31.2% of participants' self-defined area fell inside the corresponding census-tract (17).

The impact of such discrepancies in neighbourhood delimitations on residential exposure has been assessed in studies that measured access to healthcare resources (15), walking destinations (13), walkability characteristics (16) and supermarket and farmer's market (14). For example, the mean number of general practitioners, pharmacists and dentists was significantly higher in perceived neighbourhoods than in Euclidean buffer areas (of 0.42 km^2 around home) for 653 adults living in the Paris metropolitan area (France) (15). In addition, as described in a study in New Zealand, built environmental characteristics such as residential density, land use mix and connectivity that are commonly used to measure walkability score differed significantly between the five different neighbourhood limits (i.e. perceived neighbourhood, census-area, 1-mi Euclidean buffer, 1 km network buffer and activity space) (16). Altogether, results underline to what extent the use of self-defined rather than predefined neighbourhoods may lead to large differences regarding the influence of contextual factors on health-related outcomes (37).

Our study has some limitations. One is a function of the study design, based on a web survey with a web-mapping tool developed for participants to draw their neighbourhood. The difficulties that some people may face when reading a map are well recognized (38). Data are from a cross-sectional study that does not permit causal inference, and the low response rate in the SPOTLIGHT survey, at about 10% – although typical for such studies – may have resulted in selection bias as it may be that more highly motivated people participated in the survey (39). In addition, as described in Table 1, the subsamples of participants with and without self-defined neighbourhoods differed in terms of socio-demographic characteristics (with a general trend for higher age and lower socioeconomic characteristics). Caution is thus needed when generalizing our findings.

Conclusions

This study shows the feasibility of using web-based assessment (based on GIS tools) of self-defined neighbourhoods in a large-scale European study. The results indicate that size of self-defined neighbourhood varies according to both individual-level and contextual-level factors such as age, gender, educational level, length of residence, personal attachment of the neighbourhood and administrative residential neighbourhood characteristics. This has consequences for the operational definition of what a neighbourhood is and for the use of self-defined neighbourhoods as study areas when investigating associations between environmental characteristics and health outcomes including obesity.

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.12380

Table S1. Factors associated with size variation (in km^2) of self-defined neighbourhood identified by Poisson regression multilevel models after multiple imputation (N = 4374).

References

1. Papas MA, Alberg AJ, Ewing R, Helzlsouer KJ, Gary TL, Klassen AC. The built environment and obesity. *Epidemiol Rev* 2007; **29**: 129–143.
2. Lake A, Townshend T. *Alvanides S, éditeurs*. Obesogenic Environments: complexities, perception and objective meaures. United Kingdom, 2010.
3. Giskes K, van Lenthe F, Avendano-Pabon M, Brug J. A systematic review of environmental factors and obesogenic dietary intakes among adults: are we getting closer to understanding obesogenic environments? *Obes Rev* 2011; **12**: e95–106.
4. Swinburn B, Kraak V, Rutter H *et al*. Strengthening of accountability systems to create healthy food environments and reduce global obesity. *The Lancet* 26 juin 2015; **385**: 2534–2545.
5. Kawachi I, Berkman LF. *Neighborhoods and Health*. Oxford University Press: New York, 2003.
6. Black JL, Macinko J. Neighborhoods and obesity. *Nutr Rev* janv 2008; **66**: 2–20.
7. Ludwig J, Sanbonmatsu L, Gennetian L *et al*. Neighborhoods, obesity, and diabetes – a randomized social experiment. *N Engl J Med* 2011; **365**: 1509–1519.
8. Chaskin RJ. Neighborhood as a unit of planning and action: a heuristic approach. *J Plan Lit* 1998; **13**: 11–30.
9. Moudon A, Lee C, Cheadle DA *et al*. Operational definitions of walkable neighborhood: theoretical and empirical insights. *Journal of Physical Activity and Health* 2006; **3**(Suppl. 1): S99–S117.
10. Schaefer-McDaniel N, Caughy MO, O'Campo P, Gearey W. Examining methodological details of neighbourhood observations and the relationship to health: a literature review. *Soc Sci Med 1982* 2010; **70**: 277–292.
11. Coulton CJ, Korbin J, Chan T, Su M. Mapping residents' perceptions of neighborhood boundaries: a methodological note. *Am J Community Psychol* 2001; **29**: 371–383.
12. Coulton CJ, Jennings MZ, Chan T. How big is my neighborhood? Individual and contextual effects on perceptions of neighborhood scale. *Am J Community Psychol* 2013; **51**: 140–150.
13. Smith G, Gidlow C, Davey R, Foster C. What is my walking neighbourhood? A pilot study of English adults' definitions of their local walking neighbourhoods. *Int J Behav Nutr Phys Act*. 2010; **7**: 34.
14. Crawford TW, Jilcott Pitts SB, McGuirt JT, Keyserling TC, Ammerman AS. Conceptualizing and comparing neighborhood and activity space measures for food environment research. *Health Place* 2014; **30**: 215–225.
15. Vallée J, Roux GL, Chaix B, Kestens Y, Chauvin P. The 'constant size neighbourhood trap' in accessibility and health studies. *Urban Stud* 2014; **52**: 4.
16. Stewart T, Duncan S, Chaix B, Kestens Y, Schipperijn J, Schofield G. A novel assessment of adolescent mobility: a pilot study. *Int J Behav Nutr Phys Act*. 2015; **12**: 18.
17. Robinson AI, Oreskovic NM. Comparing self-identified and census-defined neighborhoods among adolescents using GPS and accelerometer. *Int J Health Geogr* 2013; **12**: 57.
18. Haney WG, Knowles ES. Perception of neighborhoods by city and suburban residents. *Hum Ecol* 1978; **6**: 201–214.
19. Pebley AR, Sastry N. Our place: perceived neighborhood size and names in Los Angeles. University of California–Los Angeles: California Center for Population Research; 2009 p. 39. Report No.: CCPR-2009-026.
20. Sastry N, Pebley AR, Zonta M. Neighborhood definitions and the spatial dimension of daily life in Los Angeles. *Calif Cent Popul Res*, UCLA; 2002. Available at: https://www.rand.org/content/dam/rand/pubs/drafts/2006/DRU2400.8.pdf.
21. Guest AM, Lee BA. How urbanites define their neighborhoods. *Popul Environ* 1984; **7**: 32–56.
22. 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.
23. 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**: 10.
24. European Environment Agency. Towards an urban atlas: assessment of spatial data on 25 European cities and urban areas. 2002.
25. Dillman DA. *Mail and internet surveys: the tailored design method*. John Wiley Co.: New York, 2007, pp. 552.

26. 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.
27. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley and Sons: New York, NY, 1987.
28. Bodner TE. What improves with increased missing data imputations? *Struct Equ Model* 2008; **15**: 651–675.
29. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011; **30**: 377–399.
30. Glass TA, Balfour JL. Neighborhoods, aging, and functional limitations. In: Kawachi I, Berkman LF (eds). *Neighborhoods and Health*. Oxford University Press: New York, 2003, pp. 303–334.
31. Lord S, Després C, Ramadier T. When mobility makes sense: a qualitative and longitudinal study of the daily mobility of the elderly. *J Environ Psychol* 2011; **31**: 52–61.
32. Nader-Hallier B. Les frontières du territoire en ville. *Rev Gériatrie*. 2013; **3**: 171–181.
33. Vallée J, Chauvin P. Investigating the effects of medical density on health-seeking behaviours using a multiscale approach to residential and activity spaces: results from a prospective cohort study in the Paris metropolitan area, France. *Int J Health Geogr*. 2012; **11**: 54.
34. Dias P, Ramadier T. Social trajectory and socio-spatial representation of urban space: the relation between social and cognitive structures. *J Environ Psychol*. 2015; **41**: 135–144.
35. Keene D, Bader M, Ailshire J. Length of residence and social integration: the contingent effects of neighborhood poverty. *Health Place*. 2013; **21**: 171–178.
36. Kearns A, Parkinson M. The significance of neighbourhood. *Urban Stud* 2001; **38**: 2103–2110.
37. Vallée J, Shareck M. Re: « examination of how neighborhood definition influences measurements of youths » access to tobacco retailers: a methodological note on spatial misclassification. *Am J Epidemiol* 2014 kwt436.
38. Lloyd RE, Bunch RL. Individual differences in map reading spatial abilities using perceptual and memory processes. *Cartogr Geogr Inf Sci* 2005; **32**: 33–46.
39. Klesges RC, Eck LH, Mellon MW, Fulliton W, Somes GW, Hanson CL. The accuracy of self-reports of physical activity. *Med Sci Sports Exerc. oct* 1990; **22**: 690–697.