BMJ Open

Contextual characteristics associated with the perceived neighborhood scale in a large urban center in Brazil

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-021445
Article Type:	Research
Date Submitted by the Author:	05-Jan-2018
Complete List of Authors:	de Almeida Celio, Fabiano; Universidade Federal de Minas Gerais, Public health Friche, Amélia Augusta ; Universidade Federal de Minas Gerais Faculdade de Medicina, Public health Jennings, M. Zane; Case Western Reserve University Jack Joseph and Morton Mandel School of Applied Social Sciences Andrade, Amanda Cristina; Universidade Federal de Minas Gerais Faculdade de Medicina, public health Xavier, Cesar ; faculdade de saúde e ecologia humana Proietti, Fernando; Faculdade de Saúde e Ecologia Humana; Centro de Pesquisas Rene Rachou Coulton, Claudia J.; Case Western Reserve University Jack Joseph and Morton Mandel School of Applied Social Sciences Caiaffa, Waleska; Federal University of Minas Gerais, Medicine
Keywords:	neighborhood, urban health, Perceived Neighborhood, neighborhood scale, Community perceptions

Contextual characteristics associated with the perceived neighborhood scale in a large urban center in Brazil

Corresponding author: Fabiano de Almeida Célio, maitiz@gmail.com**, +55 31 99877 7761**

1) **Fabiano de Almeida Célio**[1], maitiz@gmail.com; 2)**Amélia Augusta de Lima Friche**[1], gutafriche@gmail.com; 3) **M. Zane Jennings**[2], mzj@case.edu; 4) **Amanda Cristina de Souza Andrade**[1],amandasouza_est@yahoo.com.br; 5) **Cesar Coelho Xavier**[3], cesarcxavier@gmail.com; 6) **Fernando Augusto Proietti**[3], fernandoaproietti@gmail.com; 7) **Claudia J. Coulton**[2], claudia.coulton@case.edu; 8) **Waleska Teixeira Caiaffa**[1], caiaffa.waleska@gmail.com

1. Belo Horizonte Observatory for Urban Health, School of Medicine, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil,
2.Case Western Reserve University Jack Joseph and Morton Mandel School of Applied Social Sciences, Cleveland OH, USA
3. Faculdade de Saúde e Ecologia Humana, Vespasiano, Minas Gerais, Brazil
4. Centro de Pesquisas Rene Rachou, Belo Horizonte, Minas Gerais, Brazil

Keywords: neighborhood, urban health; Perceived Neighborhood; neighborhood scale, Community perceptions

Word count: 3628

Reference count: 48

Abstract

Introduction: Health outcomes have been associated with the physical and social characteristics of the neighborhood, but still there is no enough information about the contextual factors related with perceived neighborhood scale**. Objective:** To identify the contextual factors associated with the self-perceived neighborhood scale. **Methods:** We analyzed data from a multistage household survey conducted in Belo Horizonte, Brazil, 2008-2009. The dependent variable was perceived neighborhood encoded as an ordinal scale based upon a brief description of the concept of the neighborhood and two independent scales relating distance, expressed in geographical and time approach. Street connectivity, demographic density and residents' perceptions of the neighborhoods' physical and social environment were used as contextual predictors. Individual variables were used for adjustments. Multilevel ordinal logistic regression models estimated the association between perceived neighborhood scale and contextual characteristics. **Results:** Better perception of walking environment (OR=2.96; CI 95%: 1.29 – 3.82) and violence (OR=1.35; CI95%: 1.12 – 1.62) was associated with perceptions of larger neighborhood, even adjust by individual's

characteristics. **Conclusion:** There are contextual factors associated with self-perceived neighborhood scale. Careful definition of neighborhood scale is a key factor in improving the results of eco-epidemiological studies. Although these findings must be further explored in other studies, these results can contribute for the debate of a better understanding of an appropriate choice of neighborhood scale, especially for cities in Latin America.

Strengths and limitations of this study

- Large sample of an urban center in Latin America
- Analysis includes individual and contextual factors
- Neighborhood definition that can be obtained by closed question
- Analysis was able to identify the context factors associated with the perceived neighborhood scale
- Analysis takes into account physical and social factors of the neighborhood

Introduction

Eco-epidemiological research has increasingly used the concept of neighborhood as the geographical area within which physical and social environmental features affect individuals' health outcomes, as part of an emphasis on a more holistic understanding of the factors and processes shaping health outcomes within urban areas [1].

Features in the neighborhood help explain inequalities in health; can be used in studies aiming to evaluate community interventions that intended to improve health outcomes [2-4] and, have been shown to be predictive of health outcomes and health-affecting behaviors such as cardiovascular diseases [5], sexually transmitted diseases [2], mental illness [6], physical activity [7 8], among others [9-12].

However, neighborhood is a complex concept and its definitions in epidemiological studies vary widely [13], with different methodological approaches [14]. Chaix et al (2009) describe two approaches that have been used to define neighborhood in epidemiological research: *the territorial neighborhood and the ego-centered neighborhood approaches.*

Territorial neighborhoods, drawn from ecological and social perspectives, are made up of the complex social, economic, and historical relationships between their residents, local geography, and the larger urban structure around them, and have an emergent identity apart from the perceptions of individuals living within them. Researchers using this approach often select administratively-defined, mutually exclusive geographic units such as census units or municipal boundaries as a proxy for neighborhood [13 15]. Assuming resident homogeneity [16 17], this approach is adopted because secondary data is often easily available and spatial references are obtainable, which facilitates reproducibility and comparability across studies or over time. However, territorial neighborhoods consider same areas for different individuals and due to this, individual differences in neighborhood experience and exposure cannot be captured under this approach [18 19]. When the same area is attributed to several individuals in a given region, the potential for error is introduced because individuals are not exposed in a homogeneous way to the physical and social environment of the territory.

The second approach is called *ego-centered neighborhood*s and is drawn upon the idea that the contextual factors affecting individuals will differ depending on the actual location and particular geographic circumstances of those individuals. Several techniques can be used to define this approach. Most importantly, the ego-centered neighborhood results in neighborhoods that may overlap, are not mutually exclusive, and which are specific to the

household or individual resident [13]. This approach can be operationalized in three different ways. One uses a buffer, generally a circular area centered upon the individual's residence, resulting in neighborhoods of the same size, typically, though made up of different areas, which may overlap one another but are not identical. The second approach involves using individual behavioral activity spaces measured by GPS. This captures each individual's movements and activities, creating a unique measure of contextual exposure [9 20]. The third one uses individually perceived neighborhoods.

Perceived neighborhoods, in turn, can be identify by different strategies. Residents may be asked to identify or draw their neighborhood on a map [21-24] or alternatively, researchers may ask residents how large they consider their neighborhood to be or how long it takes to walk from the resident's house to the end of their neighborhood [13 25-27]. This last technique has the advantage of being easily understood by residents and quickly and inexpensively done by researchers.

Despite the methods, neighborhood scale needs to be carefully set. When it is not correctly operationalized and defined, the measures derived can be considered problematic and questionable. Consequently, the understanding of health impacts through the lens of the neighborhood can be impoverished [28]. One problem that may arise is known in geography as the [29] "modifiable area unit problem". By aggregating epidemiological data in differently sized territorial units, different exposure measures can be found and consequently different results can be obtained between studies, making it difficult or even impossible to compare findings. Generally, the error of choice of territorial unit of analysis is non-differential, which may underestimate association measures or even not find associations when they do exist [30].

The attributes that make the neighborhood of an individual a singular place are commonly described as: a) social interaction; b) social norms and collective effectiveness; c) institutional resources (schools, health facilities and others) and d) routine activities within the neighborhood. As we can see, it is difficult not to incur some kind of neighborhood boundary definition error when the internal dynamics of the place under study are unknown [31].

Perceived neighborhood scale has been found to be related to individual characteristics such as socioeconomic position, employment, evaluation of the aesthetic aspects, the number of relatives living in the same neighborhood and familiarity with many people in the neighborhood [27]. However, the scale of perceived neighborhood can be influenced by

contextual factors as population density, land use patterns, and collective efficacy [21]. The connectivity of the streets that directly influence the number of routes available to the various points of interest within a neighborhood can also influence the perception of its size, because connectivity may change the way residents use and circulate in physical space [32].

This work, by investigating perceived neighborhood scale, addresses an important methodological question, which concerns the appropriate scale of territorial units of analysis, reducing possible errors inherent to the process of investigating neighborhood impact on health outcomes. Despite research results indicating a relationship between neighborhood and health, it is still rare to find studies that measure the influence of contextual factors as shaped by perceived neighborhood scale. In Latin America, we have not found any studies with this same purpose. Therefore, the objective of this study is to analyze the context attributes associated with the perceived neighborhood scale in a large urban center in Brazil.

Methods

Data and Sample

The data for this study come from a cross-sectional population-based study called *BH Health Study*, conducted by the Belo Horizonte Observatory for Urban Health (OSUBH), nested at the Federal University of Minas Gerais (UFMG), in 2008-9. The participants of the study were residents belonging to two of the nine sanitary districts of the Belo Horizonte City: Barreiro and West. These districts were selected because they presented heterogeneity within the city in relation to social, socio-demographic and health indicators [33-35].

A stratified sample was selected by a three-stage process. To ensure the presence of residents in all socio-economic levels, the study area was divided according to the health vulnerability index [36], a geocoded index created by combining social, demographic, economic, and health indicators from different sources for each census tract. At the end of first and second steps of sampling process, 149 census tracts and 4,048 households were selected. In the third stage, a resident over 18 years old was randomly selected in each of the identified households [37].

Contextual predictors

The dependent variable for this study is the perceived neighborhood scale that was encoded as an ordinal variable originally with 7 options. To obtain the scale the interviewer read a brief description of the concept of the neighborhood: "neighborhood is the place where you live

and perform routine tasks such as going to the bakery, grocery store, local businesses, visit your neighbors and walk. Neighborhood can be understood as the area where you recognize most of the people". Then the interviewee was asked: "Thinking of your neighborhood, would you describe it as including: (1) The next-door houses; (2) The block or street you live on; (3) within 5 blocks; (4) within ten blocks; (5) more than ten blocks away, but less than your neighborhood; (6) your neighborhood; (7) your neighborhood and neighborhood nearby. Subsequently, this variable was recoded, using as reference an additional measure of neighborhood scale. This was a continuous variable obtained from the following question: "How much time in minutes would you spend walking from the door of your house to the end of what you consider your neighborhood?". The mean of walking time obtained for each of the seven options of the first ordinal question variable was used to collapse the final dependent variable into four options. This procedure was adopted by considering the non-overlapping of the 95% confidence intervals (95% CI) between each stratum. Thus, the outcome variable named Perceived neighborhood scale was recoded into four categories: (1) up to the block or street you live on; (2) within 5 blocks; (3) within ten blocks; and (4) more than ten blocks away.

Independent variables

The independent variables were chosen based on the theoretical (Figure 1) model using other studies [4 27]. The variables relating to the physical and social environment of the neighborhood were obtained from scales created by Friche et al [38]. Aggregated for each census tract the scales provide a continuous score ranging from 1 to 4. In this study, we used the following scales: aesthetic quality, walking environment, and violence.

The aesthetic quality scale was obtained by asking the participants the following about their neighborhood: 1) Has trash a litter on the streets and sidewalks? 2) Is pleasant for children? 3) Is pleasant for young children and adolescents? 4) Has trees that make the environment pleasant?

The walking environment scale was obtained by asking the participants the following about their neighborhood: 1) How you evaluate the public places for sports and leisure? 2) How you evaluate the traffic? 3) Are there stores at a distance you can walk? 4) Is it easy to walk? 5) How often do you see other people walking? 6. How often do you see other people exercising? 7) Do you feel safe walking during the day?

The violence scale was comprised by the following questions: During the past 12 months, did

you see or heard about: 1) People being mugged in the neighborhood streets; 2) People fighting using weapon; 3) People being killed by guns; 4) People being victims of sexual violence; 5) Women of the neighborhood being beaten by their husband and/or partners or relatives; 6) Children or adolescents of the neighborhood assaulted or victims of violence by their parents.

This study also used contextual variables from census tract and collected by the city hall to administrative purpose. Street connectivity drawn from all street segments of the area of the study was obtained using Dephmapx software. This software handles the street segment as if it were an axial line, and quantifies the segments that intersect each of this lines [32]. The software delivers a score between 0 a 9, where 0 represents streets with low connectivity and 9 represents high connected streets[39]. The final variable was skewed, with low prevalence of extreme values, so it was recoded into three categories: low connectivity (0 to 3); medium connectivity (4) and, high connectivity (5 to 9).

Population density was calculated for each census tract using data from the 2010 National Census [40].

Individual variables

Individual variables were included for adjustments due to being found as predictors of neighborhood scale in previous studies [21 27]. They were: gender, age (in years), employment status, length of residence in home (in years), presence of children under 10 years of age in household, number of relatives in the same neighborhood (none to all), number of people who pass in front of participant's house that are known to them (none to all), and a composite indicator, named national economic index (NEI), which depicts the current socioeconomic position of the individual, [41], based upon consumer goods instead of income.

Statistical analyses

A descriptive analysis was carried out followed by analysis of association between size of the neighborhood scale and contextual features estimated by a multilevel ordinal logistic regression model. The first level consisted of the individual-level and the second consisted of the neighborhood-level variables.

A fixed effects model with random interceptors with a logit function was used to estimate the odds ratio (OR) and the confidence interval (CI 95%) [42]. The median values of the odds

ratio (MOR) and the percentage of variance reduction were calculated. The Akaike Information Criterion (AIC) was used to compare models, selecting the best model as the one with lowest AIC [43].

First, a null model (only the random intercept) was estimated to assess the contextual effect and then a univariate analysis was performed with a multilevel ordinal logistic regression for each of the contextual variables. Second, those independent variables with a coefficient that was significant at $p \leq 0.20$ in the univariate were included as level 2 variables in the multiple analysis. Finally, we added the individual characteristics (age, gender, employment status, number of parents and friend in the neighborhood, recognize people passing by the door of your house, residence length in the same neighborhood, presence of child younger than 10 and socioeconomic position) at level 1 for adjustment.

The analyses were performed in the software STATA (Stata Corp., College Station, Texas), version 12.0. For all analysis we used the svy command [44], that considers the complex design and sampling weights. For all models, we considered a significance level of 5%.

Ethical issues

The study was approved by the Research Ethics Committee of the UFMG through opinion ETIC n° 253/06. All participants signed informed consent.

Results

We found that 57.8% of the participants considered their neighborhood to be their own house until the end of the block, 23.3% considered their neighborhood to be the 5 closest blocks, 7.4% the nearest 10 blocks to their residence and 11.5% considered their neighborhood to be larger than 10 blocks from their home.

There was a linear relationship between the size of the perceived neighborhood and the time to walk to his/her neighborhood end, with the following average times in minutes for each neighborhood size stratum: 6.1, 13.5, 19.8 and 29.2.

The distribution of street connectivity was almost the same with 39.7% of streets with connectivity between 0 and 3; 24.8% with value 4 and 35.4% with values between 5 to 9. The mean population density was 12,264 residents/km^2 (685.9). All these results are on table 1.

The multi-level model analysis began with the null model. The perceived neighborhood size had significant variation within census tract, given the likelihood-ratio test ($p < 0.001$). The analysis showed the following scales and variables significantly associated with the self-perceived neighborhood size (table 2): walking environment (OR=2.96; CI 95%: 1.29 – 3.82), violence (OR=1.35; CI95%: 1.12 – 1.62), female gender (OR= 0.81; CI 95%: 0.68 – 0.96), greater number of relatives living in the neighborhood (OR= 4.63; CI 95%: 2.84 – 7.57), recognition of more people in the neighborhood (OR= 3.33; CI 95%: 1.72 – 6.25) and, socioeconomic position (NEI) (OR= 1.17; CI 95%: 1.06 – 1.29) (Table 3).

Based on the Akaike Information Criteria (AIC) the best models were adjusted by individual variables at level 1. The median value of the odds ratio comes from the median value between the area with highest odds to have more larger perceived neighborhood scale and the area with lowest odds when randomly picking out two areas [42]. We found a MOR of 1.62 for the null model; 1.58 for the model with the contextual predictors and also 1.58 for the one adjusted by the individual variables. The results of proportional change in variance show that the contextual predictors explained 10.7% of the total variance and the model with contextual and individual variables explained 9.0%. Although the best model based on the AIC criterion is the context plus individual variables, the model only with contextual variables has more variation when compared to the null model. It means that 10.0% of the contextual variance of perceived neighborhood scale was attributed to the contextual factors and that when we added the individual level it decreased a little to 9.0% (table 3).

Discussion

Perceived neighborhood scale was relatively small for many residents: 57.8% of the participants considered their neighborhood as residences closest to their home until the end of the block. Also, contextual factors such as perceived good environment conditions for walking and indicators of a violent environment were associated with a larger perceived neighborhood scale, even adjusted by individual-level variables.

These findings allow some comparison with previous studies, but care must be taken because each study has different approaches in measuring perceived neighborhood scale. A study carried out in the city of Los Angeles [45] reported that 35.1% of the interviewees considered their neighborhood to be a block or street that they live on, 25.0% several blocks or streets in each direction, 28.8% as an area within a 15-minute walk and 13.1% considered their neighborhood as an area larger than a 15-minute walk. In other words, 86.9% of participants

considered their neighborhood as an area larger than 15-minute walk. This result is very similar to what we found, when we look to the average time taken to leave the self-perceived neighborhood within each stratum, 81.1% of participants considered their neighborhood as an area smaller than a 15-minute walk.

Another study conducted in different areas of Seattle [25] found that 46.4% of participants considered their neighborhood as their own residential unit to no more than one block in each direction.

However, studies using maps as an approach to measure the perceived neighborhood scale found much larger neighborhoods. A study [24] conducted in five Europe urban regions found a mean of 1.96 km^2. Similar results were found in a small study conducted with adolescents in Boston, where it was reported a mean area of 1.82 km^2. Stewart et al [46], found in a pilot study conducted in Auckland (New Zealand) an area of 3.54 km^2; Coulton et al [21], in a study with 6,224 adults in low-income communities in 10 US cities found an area of 2.33 km^2. A study conducted in Helsinki and Espoo, Finland, with 15,982 persons, that calculated the area inside the most visited points in a neighborhood, found an average area of 1.07 km^2 [47].

Despite the heterogeneities in the sampling and methods utilized between studies, researches that use maps found larger neighborhoods, indicating a possible relationship to the methodology used to access the perceived neighborhood scale. A possible explanation is that it is easier to remember important points in neighborhoods when participants look at a map. Using an open- or closed-ended question does not provide that kind of specific context.

The results of the multilevel model show us that there are contextual factors associated with perceived neighborhood scale. We found associations with the perceptions of the walking environment and with violence. The interpretation of the results of the scales should be done based on the analysis of the behavior of its score [38]. The walking environment scale had highest values when the census tract had more people who reported that their neighborhoods have a physical environmental that encourage mobility and external activities. To our knowledge, the literature does not report a similar relationship, but it is plausible that an area that stimulates the mobility of people, facilitating diverse activities within the neighborhood could be also related with a large perception of neighborhood scale.

Violence scale, which reports higher values for more violence neighborhoods, was associated with larger perceived neighborhood scale. This appears contradictory at first glance, but

people with larger perceived neighborhoods are likely to have greater social contact and exposure to the environment and may therefore be able to identify the problems within the neighborhood.

Regarding connectivity, we found association between high street connectivity and larger perceived neighborhood scale only in a univariate analysis, despite a negative association found in another publication[21]. After adjustments connectivity was no longer significant although the plausibility of the association; highly connected streets tend to be in busier places with a high demographic density and intense automobile traffic that hampers social contact and favors less extensive perceptions.

Demographic density was not associated with perceived neighborhood scale. The literature consulted differs in relation to this variable. Some studies have found [21 23] association between smaller perceived neighborhood scale and greater population density. Others have reported an association between higher population densities and larger neighborhoods [24 48] and further one, such ours, found no relationship [26 45]. However, neighborhoods with a high population density, especially if car traffic is intense could also impoverished social contact among neighbors favoring a lower neighborhood perception, in the same direction of connectivity.

This study has specific limitations that need to be mentioned as listed: 1. The use of a closed-ended question to obtain the perceived neighborhood scale does not specify the spaces that individuals are actually exposed to; 2. The cross-sectional design of the study limits the interpretation of some results, due to the possibility of reverse causality; 3. The results of this study are from a large urban center and are not necessarily valid for smaller cities and rural areas; 4. Perception of the neighborhood scale may differ depending upon the age of the participants since those younger than 18 years were not included in this study.

The identification of the contextual factors associated with the perception of neighborhood scale have important methodological implications, especially for studies that intend to investigate an association of social factors of the neighborhood with health events. The perceived neighborhood scale is a fundamental tool for the creation of more precise and coherent neighborhood boundaries informed by the places actually experienced by individuals.

One of the motivations of this study is related with the fact of a large amount of research in eco-epidemiology and community practice tends to use artificial definitions of neighborhoods

boundaries. The results of this study show us that we have a big heterogeneity on perceived neighborhood scale, reinforcing the argument that researchers need to use more personalized way to define neighborhood boundaries. Most of researches use census tracts as a proxy of neighborhood due the availability of data aggregated to this level, but the increase use of GIS technique has been supported a more individualized neighborhood definition that can be used to avoid problems regarding the choice of neighborhood size and its operationalization. A better neighborhood definition will help future eco-epidemiological researches that can delivered a more robust evidence to support a better community practices.

References

1. Roux AVD. Health in cities: is a systems approach needed? Cadernos de Saúde Pública 2015;**31**:9-13
2. Cohen D, Spear S, Scribner R, et al. " Broken windows" and the risk of gonorrhea. American journal of public health 2000;**90**(2):230
3. Schaefer-McDaniel N, Caughy MOB, O'Campo P, et al. Examining methodological details of neighbourhood observations and the relationship to health: a literature review. Social science & medicine 2010;**70**(2):277-92
4. Diez Roux AV, Mair C. Neighborhoods and health. Annals of the New York Academy of Sciences 2010;**1186**(1):125-45
5. Roux AVD, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. New England Journal of Medicine 2001;**345**(2):99-106
6. Mair CF, Roux AVD, Galea S. Are neighborhood characteristics associated with depressive symptoms? A critical review. Journal of epidemiology and community health 2008:jech. 2007.066605
7. Andrade ACdS, Peixoto SV, Friche AAdL, et al. Social context of neighborhood and socioeconomic status on leisure-time physical activity in a Brazilian urban center: The BH Health Study. Cadernos de Saúde Pública 2015;**31**:136-47
8. Fernandes AP, Andrade ACdS, Ramos CGC, et al. Leisure-time physical activity in the vicinity of Academias da Cidade Program in Belo Horizonte, Minas Gerais State, Brazil: the impact of a health promotion program on the community. Cadernos de Saúde Pública 2015;**31**:195-207
9. Chaix B, Kestens Y, Bean K, et al. Cohort Profile: Residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases—The RECORD Cohort Study. International Journal of Epidemiology 2011;**107**:1-10
10. Ludwig J, Duncan GJ, Gennetian LA, et al. Neighborhood effects on the long-term well-being of low-income adults. Science 2012;**337**(6101):1505-10
11. Kaiser P, Diez Roux AV, Mujahid M, et al. Neighborhood environments and incident hypertension in the multi-ethnic study of atherosclerosis. American journal of epidemiology 2016;**183**(11):988-97
12. Johnson DA, Simonelli G, Moore K, et al. The neighborhood social environment and objective measures of sleep in the Multi-Ethnic Study of Atherosclerosis. Sleep 2017;**40**(1)
13. Chaix B, Merlo J, Evans D, et al. Neighbourhoods in eco-epidemiologic research: delimiting personal exposure areas. A response to Riva, Gauvin, Apparicio and Brodeur. Social science & medicine 2009;**69**(9):1306-10
14. Flowerdew R, Manley DJ, Sabel CE. Neighbourhood effects on health: does it matter where you draw the boundaries? Social science & medicine 2008;**66**(6):1241-55
15. Galster G. On the nature of neighbourhood. Urban studies 2001;**38**(12):2111-24
16. Santos SM, Chor D, Werneck GL. Demarcation of local neighborhoods to study relations between contextual factors and health. International journal of health geographics 2010;**9**(1):1
17. Clapp JM, Wang Y. Defining neighborhood boundaries: Are census tracts obsolete? Journal of Urban Economics 2006;**59**(2):259-84
18. Chaskin RJ. Perspectives on neighborhood and community: a review of the literature. The Social Service Review 1997:521-47
19. Chappell NL, Funk LM, Allan D. Defining community boundaries in health promotion research. American Journal of Health Promotion 2006;**21**(2):119-26
20. Chaix B, Meline J, Duncan S, et al. GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? Health & Place 2013;**21**:46-51
21. Coulton C, Jennings MZ, Chan T. How Big is My Neighborhood? Individual and Contextual Effects

on Perceptions of Neighborhood Scale. American Journal of Community Psychology 2012:1-11 doi: 10.1007/s10464-012-9550-6[published Online First: Epub Date]|.

22. Coulton CJ, Korbin J, Chan T, et al. Mapping residents' perceptions of neighborhood boundaries: a methodological note. American journal of community psychology 2001;**29**(2):371-83
23. Haney W, Knowles E. Perception of neighborhoods by city and suburban residents. Human Ecology 1978;**6**(2):201-14 doi: 10.1007/bf00889095[published Online First: Epub Date]|.
24. Charreire H, Feuillet T, Roda C, et al. Self-defined residential neighbourhoods: size variations and correlates across five European urban regions. Obesity Reviews 2016;**17**(S1):9-18
25. Guest AM, Lee BA. How urbanites define their neighborhoods. Population and Environment 1984;**7**(1):32-56
26. Pebley AR, Sastry N. Our place: perceived neighborhood size and names in Los Angeles. Los Angeles: California Center for Population Research, University of California 2009
27. Célio FdA, Xavier CC, Andrade ACdS, et al. Características individuais associadas à autopercepção da extensão territorial da vizinhança. Cadernos de Saúde Pública 2014;**30**:1935-46
28. Cutchin MP, Eschbach K, Mair CA, et al. The socio-spatial neighborhood estimation method: an approach to operationalizing the neighborhood concept. Health & place 2011;**17**(5):1113-21
29. Fotheringham AS, Wong DW. The modifiable areal unit problem in multivariate statistical analysis. Environment and planning A 1991;**23**(7):1025-44
30. Duncan DT, Kawachi I, Subramanian SV, et al. Examination of How Neighborhood Definition Influences Measurements of Youths' Access to Tobacco Retailers: A Methodological Note on Spatial Misclassification. American Journal of Epidemiology 2013 doi: 10.1093/aje/kwt251[published Online First: Epub Date]|.
31. Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing" neighborhood effects": Social processes and new directions in research. Annual review of sociology 2002:443-78
32. Hillier B. Spatial sustainability in cities: organic patterns and sustainable forms. 2009
33. Andrade RGd, Chaves OC, Costa DAdS, et al. Overweight in men and women among urban area residents: individual factors and socioeconomic context. Cadernos de Saúde Pública 2015;**31**:148-58
34. Camargos VP, César CC, Caiaffa WT, et al. Imputação múltipla e análise de casos completos em modelos de regressão logística: uma avaliação prática do impacto das perdas em covariáveis. Cad Saúde Pública 2011;**27**:2299-313
35. Ferreira AD, César CC, Malta DC, et al. Validade de estimativas obtidas por inquérito telefônico: comparação entre VIGITEL 2008 e Inquérito Saúde em Beagá. Rev Bras Epidemiol 2011;**14**(1):16-30
36. Gerência de Epidemiologia e Informação Secretaria Municipal de Saúde de Belo Horizonte. Índice de vulnerabilidade à saúde. Secondary Índice de vulnerabilidade à saúde 2003. http://www.pbh.gov.br/smsa/biblioteca/gabinete/risco2003.
37. Friche AA, Xavier CC, Proietti FA, et al. *Saúde urbana em Belo Horizote*. 1 ed, 2015.
38. Friche AA, Diez-Roux AV, Cesar CC, et al. Assessing the Psychometric and Ecometric Properties of Neighborhood Scales in. Journal of urban health : bulletin of the New York Academy of Medicine 2012 doi: 10.1007/s11524-012-9737-z[published Online First: Epub Date]|.
39. Song Y, Knaap G-J. Measuring urban form: Is Portland winning the war on sprawl? Journal of the American Planning Association 2004;**70**(2):210-25
40. Demográfico IC. Resultados do universo. Rio de Janeiro: IBGE 2010
41. Barros AJ, Victora CG. Indicador econômico para o Brasil baseado no censo demográfico de 2000. Revista de Saúde Pública 2005;**39**(4):523-29
42. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health 2006;**60**(4):290-7 doi: 10.1136/jech.2004.029454[published Online First: Epub Date]|.
43. Twisk J. Applied multilevel analysis: a practical guide. 2006: Cambridge University Press, Cambridge, UK.

44. StataCorp L. Stata survey data reference manual, 1985.
45. Sastry N, Pebley AR, Zonta M. Neighborhood definitions and the spatial dimension of daily life in Los Angeles. California Center for Population Research 2002
46. Stewart T, Duncan S, Chaix B, et al. A novel assessment of adolescent mobility: a pilot study. International journal of behavioral nutrition and physical activity 2015;**12**(1):18
47. Hasanzadeh K, Broberg A, Kyttä M. Where is my neighborhood? A dynamic individual-based definition of home ranges and implementation of multiple evaluation criteria. Applied Geography 2017;**84**:1-10
48. Vallée J, Le Roux G, Chaix B, et al. The 'constant size neighbourhood trap'in accessibility and health studies. Urban Studies 2014:0042098014528393

Table 1. Univariate analysis of individual variables by perceived neighborhood scale: percentages, means and standard deviations.

INDIVIDUAL VARIABLES	Perceived Neighborhood scale (1-4 and %)*				OR (CI 95 %) [1]	p
	1 (57.8%)	2 (23.3%)	3 (7.4%)	4 (11.5%)		
Gender (female)	56.4	51.0	45.9	45.9	1.42 (1.20 - 1.70)	<0.001
Employment state (working)	62.0	65.6	69	73.3	1.36 (1.14 - 1.61)	<0.001
Presence of child younger than 10 years (yes)	33.1	31.5	30.6	33.0	0.95 (0.82 - 1.11)	0.540
Number of relatives and friends living in the same neighborhood (almost all)	2.92	7.07	7.83	13.03	6.30 (4.00 - 9.92)	<0.001
Recognizes most of them people passing by the door of his/her house (yes)	8.0	12.6	13.1	21.3	5.55 (3.04 - 10.11)	<0.001
	Mean (Standard deviation)				OR (CI 95 %) [1]	p
Age (years)	44.7 (0.35)	44.9 (0.57)	41.1 (0.92)	43.2 (0.74)	0.99 (0.99 - 1.00)	0.060
Socioeconomic position (NEI) **	586.8 (4.06)	601.6 (6.6)	582.8 (11.4)	601.4 (0.0)	1.13 (1.04 - 1.23)	<0.001
Time of residence in the same neighborhood (years)	14.8 (0.26)	16.8 (0.44)	16.2 (0.71)	16.6 (0.60)	1.01 (1.01 -1.02)	<0.001

* (1) up to the block or street you live on; (2) within 5 blocks; (3) within ten blocks; and (4) more than ten blocks away; OR – odds ratio; CI 95% – 95% confidence interval; NEI - national economic index; ** Odds ratio calculated based on an increase of 200 NEI points; [1] estimated by a multilevel ordinal logistic regression model, the reference category is the smaller neighborhood.

Table 2. Univariate analysis of contextual variables by perceived neighborhood scale: percentages, means and standard deviations.

CONTEXTUAL VARIABLES	NEIGHBORHOOD EXTENSION SCALE (1-4 and %)*				OR (CI 95 %) [1]	p
	1 (57.8%)	2 (23.3%)	3 (7.4%)	4 (11.5%)		
Connectivity [2]						
0 to 3	39.91	40.91	41.87	40.37	1.00	
4	23.92	23.42	36.11	25.28	1.04 (0.83 - 1.30)	0.760
5 to 9	37.17	35.63	22.02	34.34	0.85 (0.70 - 1.04)	0.120
	Mean (Standard deviation)				OR (CI 95 %) [1]	p
Aesthetic Quality scale	2.96 (0.03)	3.07 (0.04)	3.06 (0.07)	3.09 (0.04)	1.23 (1.03 - 1.46)	0.020
Walking Environment scale	3.20 (0.01)	3.24 (0.02)	3.27 (0.02)	3.28 (0.02)	3.37 (2.09 - 5.44)	<0.001
Violence scale	1.90 (0.02)	1.95 (0.03)	1.89 (0.05)	2.00 (0.04)	1.20 (1.03 - 1.40)	0.020
Safety scale	2.96 (0.03)	2.93 (0.05)	2.89 (0.09)	2.86 (0.05)	0.92 (0.79 - 1.08)	0.190
Population density (per square kilometer)	12487.35 (791.26)	11740.34 (704.83)	12274.31 (740.63)	12627.22 (865.05)	1.00 (0.99 - 1.01)	0.650

* (1) up to the block or street you live on; (2) within 5 blocks; (3) within ten blocks; and (4) more than ten blocks away; OR – odds ratio; CI 95% – 95% confidence interval. [1] estimated by a multilevel ordinal logistic regression model, the reference category is the smaller neighborhood; [2] 0 means poorly connected streets and 9 heavily connected streets.

Table 3. Multilevel ordinal logistic regression for the resident perceived neighborhood scale

Variables	Null Model	Contextual variables		Contextual variables + individual variables	
		OR (CI 95%) [1]	p	OR (CI 95%) [1]	p
Aesthetic Quality scale		1.21 (0.97 - 1.41)	0.060	1.13 (0.92 -1.39)	0.230
Walking Environment scale		2.96 (1.71 - 5.13)	<0.001	2.22 (1.29 -3.82)	<0.001
Violence scale		1.35 (1.12 - 1.62)	<0.001	1.23 (1.01 - 1.51)	0.040
Safety scale		0.97 (0.82 - 1.14)	0.710	0.99 (0.83 -1.19)	0.950
Connectivity [2]					
4		1.00 (0.80 -1.26)	0.940	1.06 (0.85 - 1.34)	0.570
5 e 9		0.82 (0.67 - 1.01)	0.060	0.89 (0.72 - 1.11)	0.310
INDIVIDUAL					
Age (years)				1.00 (0.99 - 1.00)	0.350
Gender (female)				0.81 (0.68 - 0.96)	0.020
Number of relatives and friends living in the same neighborhood (almost all)				4.63 (2.84 - 7.57)	<0.001
Recognizes most of them people passing by the door of his/her house (yes)				3.33 (1.72 - 6.25)	<0.001
Employment state (working)				1.26 (1.06 - 1.50)	0.010
Time of residence in the same neighborhood (years)				1.01 (1.00 - 1.01)	0.130
Presence of child younger than 10 years old (yes)				0.97 (0.81 - 1.16)	0.740
Socioeconomic position**				1.17 (1.06 - 1.29)	<0.001
MODEL INFORMATION					
Variance	0.2567		0.2292		0.2336
MOR	1.62		1.58		1.58
Proportional change in variance	-		10.71		9.00
AIC	8749.26		8668.44		8091.83

OR – odds ratio; CI 95% – 95% confidence interval; NEI - national economic index; MOR - median values of the odds ratio; AIC - Akaike Information Criterion; ** Odds ratio calculated based on an increase of 200 NEI points; [1] estimated by a multilevel ordinal logistic regression model, the reference category is the smaller neighborhood; [2] 0 means connected streets and 9 heavily connected streets.

Figure legends

Figure 1 - Theoretical model for perceived neighborhood scale

Contributors:

de Almeida Celio, Fabiano de Almeida: Fabiano participated in the design of the study, its conception. Fabiano was responsible for writing the article and the following activities: He setting up the database, he made the statistical analysis, he reviewed the text, and approved the final version.

Friche, Amélia Augusta de Lima: Professor Friche participated in the design and construction of the field work; reviewed all versions of the work, ensured its accuracy and integrity and she approved the final version.

Jennings, M. Zane: Zane participated in design; ensured the accuracy and integrity of the data; approved all versions of the work and the final version.

Andrade, Amanda Cristina de Souza: Amanda participated in the design of the work, review all versions, statistical analysis, ensured the accuracy and precision of the data, approved the final version.

Xavier, Cesar Coelho: Professor Xavier participated in the design and construction of the field work; reviewed all versions of the work, ensured its accuracy and integrity and he approved the final version.

Proietti, Fernando: Professor Proietti participated in the design and construction of the field work; reviewed all versions of the work, ensured its accuracy and integrity and he approved the final version.

Coulton, Claudia J.: Professor Coulton participated in design; ensured the accuracy and integrity of the data; approved all versions of the work and the final version.

Caiaffa, Waleska Teixeira: Professor Caiaffa participated in the design and construction of the field work; reviewed all versions of the work, ensured its accuracy and integrity and she approved the final version.

Funding by:

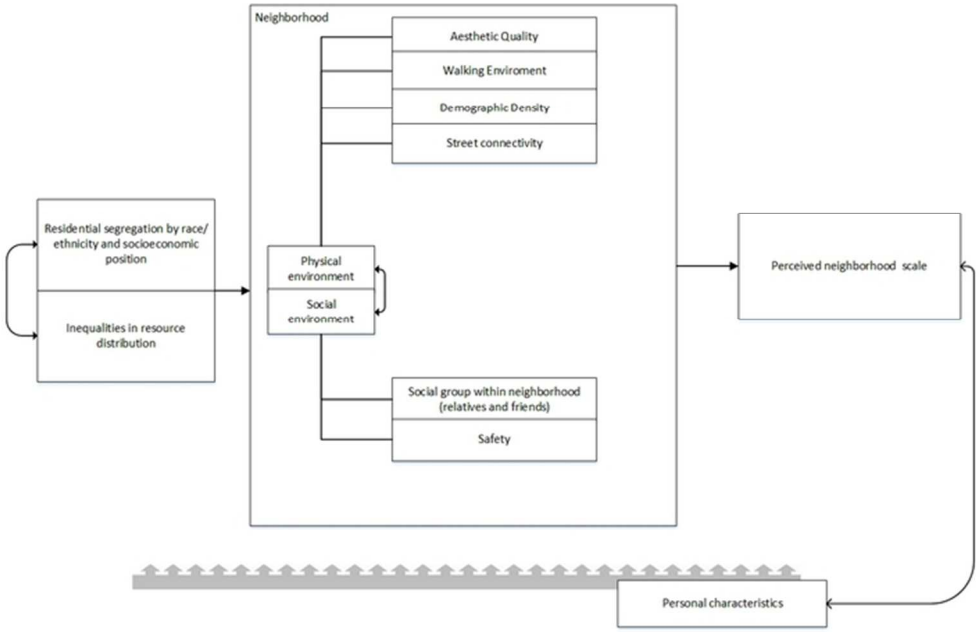
National Health Fund of the Ministry of Health, Fapemig, CNPq, NIH / Fogarty International Center.

Competing Interests Declaration:

All authors declare they have no actual or potential competing financial interest.

Data sharing:

All relevant data are within the paper. The data underlying this study are third party data and are available to all interested researchers. To submit a proposal the OSUBH Coordinating Center at http://site.medicina.ufmg.br/osubh/contato/.



Theoretical Model for perceived neighborhood scale

30x19mm (600 x 600 DPI)

BMJ Open

Contextual characteristics associated with the perceived neighborhood scale, in a cross-sectional study in a large urban center in Brazil

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-021445.R1
Article Type:	Research
Date Submitted by the Author:	09-Apr-2018
Complete List of Authors:	de Almeida Celio, Fabiano; Universidade Federal de Minas Gerais, Public health Friche, Amélia Augusta ; Universidade Federal de Minas Gerais Faculdade de Medicina, Public health Jennings, M. Zane; Case Western Reserve University Jack Joseph and Morton Mandel School of Applied Social Sciences Andrade, Amanda Cristina; Universidade Federal de Minas Gerais Faculdade de Medicina, public health Xavier, Cesar ; faculdade de saúde e ecologia humana Proietti, Fernando; Faculdade de Saúde e Ecologia Humana; Centro de Pesquisas Rene Rachou Coulton, Claudia J.; Case Western Reserve University Jack Joseph and Morton Mandel School of Applied Social Sciences Caiaffa, Waleska; Federal University of Minas Gerais, Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Public health
Keywords:	neighborhood, urban health, Perceived Neighborhood, neighborhood scale, Community perceptions

Contextual characteristics associated with the perceived neighborhood scale, in a cross-sectional study in a large urban center in Brazil

Corresponding author: Fabiano de Almeida Célio, maitiz@gmail.com**, +55 31 99877 7761**

1) **Fabiano de Almeida Célio**[1], maitiz@gmail.com; 2)**Amélia Augusta de Lima Friche**[1], gutafriche@gmail.com; 3) **M. Zane Jennings**[2], mzj@case.edu; 4) **Amanda Cristina de Souza Andrade**[1],amandasouza_est@yahoo.com.br; 5) **Cesar Coelho Xavier**[3], cesarcxavier@gmail.com; 6) **Fernando Augusto Proietti**[3], fernandoaproietti@gmail.com; 7) **Claudia J. Coulton**[2], claudia.coulton@case.edu; 8) **Waleska Teixeira Caiaffa**[1], caiaffa.waleska@gmail.com

1. Belo Horizonte Observatory for Urban Health, School of Medicine, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil,
2.Case Western Reserve University Jack Joseph and Morton Mandel School of Applied Social Sciences, Cleveland OH, USA
3. Faculdade de Saúde e Ecologia Humana, Vespasiano, Minas Gerais, Brazil
4. Centro de Pesquisas Rene Rachou, Belo Horizonte, Minas Gerais, Brazil

Keywords: neighborhood, urban health; Perceived Neighborhood; neighborhood scale, Community perceptions

Word count: 3741

Reference count: 48

Abstract

Introduction: Health outcomes have been associated with physical and social characteristics of neighborhoods, but little is known about the relationship between contextual factors and perceived neighborhood scale**. Objective:** To identify the contextual factors associated with self-perceived neighborhood scale. **Methods:** We analyzed data from a cross-sectional population-based study in Belo Horizonte, Brazil, 2008-2009. The dependent variable was perceived neighborhood encoded as an ordinal scale based upon a brief description of the concept of the neighborhood and two independent scales relating distance, expressed in terms of geography and time. Street connectivity, demographic density and residents' perceptions of the neighborhoods' physical and social environment were used as contextual predictors. Individual characteristics were used as co-variates. Multilevel ordinal logistic regression models estimated the association between perceived neighborhood scale and contextual characteristics. **Results:** Resident's that perceive better walkability (OR=2.96; CI 95%: 1.29 – 3.82) and high amounts of violence (OR=1.35; CI95%: 1.12 – 1.62) perceived their neighborhoods to be larger, even after adjusting for individual's characteristics. **Conclusion:**

There are contextual factors that are associated with self-perceived neighborhood scale. Careful definition of neighborhood scale is a key factor in improving the results of eco-epidemiological studies. Although these findings must be further explored in other studies, these results can contribute to a better understanding of an appropriate choice of neighborhood scale, especially for cities in Latin America.

Summary Boxes:

Strengths and limitations of this study:

- Large sample of an urban center in Latin America
- Analysis includes individual and contextual factors
- Neighborhood definition that can be obtained by closed question
- Analysis was able to identify the context factors associated with the perceived neighborhood scale
- Analysis takes into account physical and social factors of the neighborhood

What is already known on this subject?

Neighborhood context has been found to affect health outcomes and health-affecting behaviors. In these studies, researchers typically select administratively defined geographic units such as census units or municipal boundaries as a proxy for neighborhood. However, studies based on census units do not take into account that there may be individual differences in how the neighborhood is experienced. Surveys that use the self-perceived neighborhood. may be advantageous in comparison to the census tract, because it is closer to the places actually experienced by the individual.

What does this study add?

Despite research results indicating a relationship between neighborhood and health, it is still rare to find studies that measure the influence of contextual factors as shaped by perceived neighborhood scale. In Latin America, we have not found any studies with this same purpose. This study investigated the contextual factors associated with perceived neighborhood scale in a large urban center in Latin America and provides an approach that may be useful to other studies that are considering a territorial unit of analysis.

Introduction

Eco-epidemiological research has increasingly used the concept of neighborhood as the geographical area within which physical and social environmental features affect individuals' health outcomes, as part of an emphasis on a more holistic understanding of the factors and processes shaping health outcomes within urban areas [1].

Features in the neighborhood help explain inequalities in health; can be used in studies aiming to evaluate community interventions that intended to improve health outcomes [2-4] and, have been shown to be predictive of health outcomes and health-affecting behaviors such as cardiovascular diseases [5], sexually transmitted diseases [2], mental illness [6], physical activity [7 8], among others [9-12].

However, neighborhood is a complex concept and its definitions in epidemiological studies vary widely [13], with different methodological approaches [14]. Chaix et al (2009) describe two approaches that have been used to define neighborhood in epidemiological research: *the territorial neighborhood and the ego-centered neighborhood approaches.*

Territorial neighborhoods are generally administrative areas corresponding to a territory-subdividing approach. But more complex definitions of territorial neighborhoods may consider built environment features and population characteristics. Researchers using this approach often select administratively-defined, mutually exclusive geographic units such as census tract or municipal boundaries as a proxy for neighborhood [13 15]. Assuming resident homogeneity [16 17], this approach is adopted because secondary data is often easily available and spatial references are obtainable, which facilitates reproducibility and comparability across studies or over time. However, territorial neighborhoods consider the same areas for different individuals and due to this, individual differences in neighborhood experience and exposure cannot be captured under this approach [18 19]. When the same area is attributed to several individuals in a given area, the potential for error is introduced because individuals may not be exposed in a homogeneous way to the physical and social environment of the territory.

The second approach is called *ego-centered neighborhood*s and is based upon the idea that the contextual factors affecting individuals will differ depending on the actual location and particular geographic circumstances of those individuals. Several techniques can be used to define this approach. Most importantly, the ego-centered neighborhood results in neighborhoods that may overlap, are not mutually exclusive, and which are specific to the

household or individual resident [13]. This approach can be operationalized in three different ways. One uses a buffer, generally a circular area centered upon the individual's residence, resulting in neighborhoods of the same size, though made up of different areas, which may overlap one another but are not identical. The second approach involves using individual behavioral activity spaces measured by GPS. This captures each individual's movements and activities, creating a unique measure of contextual exposure [9 20]. The third method relies on individuals perceived neighborhoods.

Perceived neighborhoods, in turn, can be identify by different strategies. Residents may be asked to identify or draw their neighborhood on a map [21-24] or alternatively, researchers may ask residents how large they consider their neighborhood to be or how long it takes to walk from the resident's house to the end of their neighborhood [13 25-27]. This last technique has the advantage of being easily understood by residents and quickly and inexpensively done by researchers.

Regardless of the methods, neighborhood scale needs to be carefully considered. When it is not correctly operationalized and defined, the measures derived can be considered problematic and questionable. Consequently, the understanding of health impacts through the lens of the neighborhood can be undermined [28]. One problem that may arise is known in geography as the [29] "modifiable area unit problem". Aggregating epidemiological data into differently sized territorial units can yield varying exposure measures results, making it difficult or even impossible to compare findings. Generally, the error of choice of territorial unit of analysis is non-differential, which may underestimate association measures or even not find associations when they do exist [30].

The attributes that make the neighborhood of an individual a singular place are commonly described as: a) social interaction; b) social norms and collective effectiveness; c) institutional resources (schools, health facilities and others) and d) routine activities within the neighborhood. As we can see, it is difficult not to incur some kind of neighborhood boundary definition error when the internal dynamics of the place under study are unknown [31].

Perceived neighborhood scale has been found to be related to individual characteristics such as socioeconomic position, employment, evaluation of the aesthetic aspects, the number of relatives living in the same neighborhood and familiarity with many people in the neighborhood [27]. However, the scale of perceived neighborhood can be influenced by contextual factors as population density, land use patterns, and collective efficacy [21]. The

connectivity of the streets that directly influence the number of routes available to the various points of interest within a neighborhood can also influence the perception of its size, because connectivity may change the way residents use and circulate in physical space [32].

This work, by investigating perceived neighborhood scale, addresses an important methodological question, which concerns the appropriate scale of territorial units of analysis, reducing possible errors inherent to the process of investigating neighborhood impact on health outcomes. Despite research results indicating a relationship between neighborhood and health, it is still rare to find studies that measure the influence of contextual factors as shaped by perceived neighborhood scale. In Latin America, we have not found any studies with this same purpose. Therefore, the objective of this study is to analyze the context attributes associated with the perceived neighborhood scale in a large urban center in Brazil.

Methods

Data and Sample

The data for this study come from a cross-sectional population-based study called *BH Health Study*, conducted by the Belo Horizonte Observatory for Urban Health (OSUBH), nested at the Federal University of Minas Gerais (UFMG), in 2008-9. The participants of the study were residents belonging to two of the nine sanitary districts of the Belo Horizonte City: Barreiro and West. These districts were selected because they presented heterogeneity within the city in relation to social, socio-demographic and health indicators [33-35].

A stratified sample was selected in a three-stage process. To ensure the presence of residents in all socio-economic levels, the study area was stratified by the health vulnerability index [36], a geocoded index created by combining social, demographic, economic, and health indicators from different sources for each census tract. At the end of first and second steps of sampling process, 149 census tracts and 4,048 households were randomly selected. In the third stage, one resident over 18 years old was randomly selected in each of the identified households [37].

Contextual predictors

The dependent variable for this study is the perceived neighborhood scale that was encoded as an ordinal variable originally with 7 options. To obtain the scale the interviewer read a brief description of the concept of the neighborhood: "neighborhood is the place where you live and perform routine tasks such as going to the bakery, grocery store, local businesses, visit

your neighbors and walk. Neighborhood can be understood as the area where you recognize most of the people". Then the interviewee was asked: "Thinking of your neighborhood, would you describe it as including: (1) The next-door houses; (2) The block or street you live on; (3) within 5 blocks; (4) within ten blocks; (5) more than ten blocks away, but less than your neighborhood; (6) your neighborhood; (7) your neighborhood and neighborhood nearby. Subsequently, this variable was recoded, using as reference an additional measure of neighborhood scale. This was a continuous variable obtained from the following question: "How much time in minutes would you spend walking from the door of your house to the end of what you consider your neighborhood?". The mean of walking time obtained for each of the seven options of the first ordinal question variable was used to collapse the final dependent variable into four options. This procedure was adopted by considering the non-overlapping of the 95% confidence intervals (95% CI) between each stratum. Thus, the outcome variable named Perceived neighborhood scale was recoded into four categories: (1) up to the block or street you live on; (2) within 5 blocks; (3) within ten blocks; and (4) more than ten blocks away.

Independent variables

The independent variables were chosen based on the theoretical (Figure 1) model using other studies [4 27]. The variables relating to the physical and social environment of the neighborhood were obtained from domains created by Friche et al [38]. Aggregated for each census tract the domains provide a continuous score ranging from 1 to 4. In this study, we used the following ones: aesthetic quality, walking environment, safety and violence.

The aesthetic quality domain was obtained by asking the participants the following about their neighborhood: 1) Has trash a litter on the streets and sidewalks? 2) Is pleasant for children? 3) Is pleasant for young children and adolescents? 4) Has trees that make the environment pleasant?

The walking environment domain was obtained by asking the participants the following about their neighborhood: 1) How you evaluate the public places for sports and leisure? 2) How you evaluate the traffic? 3) Are there stores at a distance you can walk? 4) Is it easy to walk? 5) How often do you see other people walking? 6. How often do you see other people exercising? 7) Do you feel safe walking during the day?

The violence domain was comprised by the following questions: During the past 12 months, did you see or heard about: 1) People being mugged in the neighborhood streets; 2) People

fighting using weapon; 3) People being killed by guns; 4) People being victims of sexual violence; 5) Women of the neighborhood being beaten by their husband and/or partners or relatives; 6) Children or adolescents of the neighborhood assaulted or victims of violence by their parents.

The safety domain was built with the following questions: In your neighborhood: 1) You feel safe walking during the night; 2) Violence is a problem

This study also used contextual variables from census tract and collected by the city hall for administrative purposes. Street connectivity drawn from all street segments of the area of the study was obtained using Dephmapx software. This software handles the street segment as if it were an axial line, and quantifies the segments that intersect each of this lines [32]. The software delivers a score between 0 a 9, where 0 represents streets with low connectivity and 9 represents high connected streets[39]. The final variable was skewed, with low prevalence of extreme values, so it was recoded into three categories: low connectivity (0 to 3); medium connectivity (4) and, high connectivity (5 to 9).

Population density was calculated for each census tract using data from the 2010 National Census [40].

Individual variables

Individual characteristics were included as co-variates that had been found to be predictors of neighborhood scale in previous studies [21 27]. They were: gender, age (in years), employment status, length of residence in home (in years), presence of children under 10 years of age in household, number of relatives in the same neighborhood (none to all), number of people who pass in front of participant's house that are known to them (none to all), and a composite indicator, named national economic index (NEI), which depicts the current socioeconomic position of the individual, [41], based upon consumer goods instead of income.

Statistical analyses

A descriptive analysis was carried out followed by analysis of association between size of the neighborhood scale and contextual features estimated by a multilevel ordinal logistic regression model. The first level consisted of the individual-level and the second consisted of the neighborhood-level variables.

A regression model with random interceptors with a logit function were used to estimate the

odds ratio (OR) and the confidence interval (CI 95%) [42]. The median values of the odds ratio (MOR) and the percentage of variance reduction were calculated. The Akaike Information Criterion (AIC) was used to compare models, selecting the best model as the one with lowest AIC [43].

First, a null model (only the random intercept) was estimated to assess the contextual effect and then a univariate analysis was performed with a multilevel ordinal logistic regression for each of the contextual variables. Second, those independent domains with a coefficient that was significant at $p \leq 0.20$ (aesthetic quality, walking environment, violence scale ad safety) in the univariate were included as level 2 variables in the multiple analysis. Finally, we added the individual characteristics (age, gender, employment status, number of parents and friend in the neighborhood, recognize people passing by the door of your house, residence length in the same neighborhood, presence of child younger than 10 and socioeconomic position) at level 1 for adjustment.

The analyses were performed in the software STATA (Stata Corp., College Station, Texas), version 12.0. For all analysis we used the svy command [44], that considers the complex design and sampling weights. For all models, we considered a significance level of 5%.

Patient and Public Involvement

The research participants were selected in two regions of the city that show great internal heterogeneity in relation to sociodemographic characteristics. The local population was previously clarified about the objectives and importance of the research. After the selection of the participant's house, the objectives of the research and the selection criteria were clarified. The eligible participant was randomly chosen inside of each house. All participants signed a free and informed consent form. The results of the study were disseminated to the population and public policy administrators of the municipality.

Ethical issues

The study was approved by the Research Ethics Committee of the UFMG through opinion ETIC n° 253/06. All participants signed informed consent.

Results

The final sample had 4,048 respondents, 53.1% were men and 46.9% were women, with ages varying between 18 and 95 years (mean = 44.4, SD = 16.9). We found that 57.8% of the participants considered their neighborhood to be their own house until the end of the block,

23.3% considered their neighborhood to be the 5 closest blocks, 7.4% the nearest 10 blocks to their residence and 11.5% considered their neighborhood to be larger than 10 blocks from their home (table 1).

There was a linear relationship between the size of the perceived neighborhood and the time to walk to his/her neighborhood end, with the following average times in minutes for each neighborhood size stratum: 6.1, 13.5, 19.8 and 29.2.

The distribution of street connectivity was almost the same with 39.7% of streets with connectivity between 0 and 3; 24.8% with value 4 and 35.4% with values between 5 to 9. The mean population density was 12,264 residents/km^2 (685.9). All these results are on table 2.

The multi-level model analysis began with the null model. The perceived neighborhood size had significant variation within census tract, given the likelihood-ratio test ($p < 0.001$). The analysis showed the following scales and variables significantly associated with the self-perceived neighborhood size (table 3): walking environment (OR=2.96; CI 95%: 1.29 – 3.82), violence (OR=1.35; CI95%: 1.12 – 1.62), female gender (OR= 0.81; CI 95%: 0.68 – 0.96), greater number of relatives living in the neighborhood (OR= 4.63; CI 95%: 2.84 – 7.57), recognition of more people in the neighborhood (OR= 3.33; CI 95%: 1.72 – 6.25) and, socioeconomic position (NEI) (OR= 1.17; CI 95%: 1.06 – 1.29).

Based on the Akaike Information Criteria (AIC) the best models were adjusted by individual variables at level 1. The median value of the odds ratio comes from the median value between the area with highest odds of a larger perceived neighborhood scale and the area with lowest odds when randomly picking out two areas [42]. We found a MOR of 1.62 for the null model; 1.58 for the model with the contextual predictors and also 1.58 for the one adjusted by the individual variables. The results of proportional change in variance show that the contextual predictors explained 10.7% of the total variance and the model with contextual and individual variables explained 9.0%. Although the best model based on the AIC criterion is the context plus individual variables, the model only with contextual variables has more variation when compared to the null model. It means that 10.0% of the contextual variance of perceived neighborhood scale was attributed to the contextual factors and that when we added the individual level it decreased a little to 9.0% (table 3).

Discussion

Perceived neighborhood scale was relatively small for many residents: 57.8% of the participants considered their neighborhood as residences closest to their home until the end of the block. Also, contextual factors such as perceived quality of environmental conditions for walking and indicators of a violent environment were associated with a larger perceived neighborhood scale, even adjusted by individual-level variables.

These findings allow some comparison with previous studies, but care must be taken because each study has different approaches in measuring perceived neighborhood. A study carried out in the city of Los Angeles [45] reported that 35.1% of the interviewees considered their neighborhood to be a block or street that they live on, 25.0% several blocks or streets in each direction, 28.8% as an area within a 15-minute walk and 13.1% considered their neighborhood as an area larger than a 15-minute walk. In other words, 86.9% of participants considered their neighborhood as an area larger than 15-minute walk. This result is very similar to what we found, when we look to the average time taken to leave the self-perceived neighborhood within each stratum, 81.1% of participants considered their neighborhood as an area smaller than a 15-minute walk.

Another study conducted in different areas of Seattle [25] found that 46.4% of participants considered their neighborhood as their own residential unit to no more than one block in each direction.

However, studies using maps as an approach to measure the perceived neighborhood found much larger neighborhoods. A study [24] conducted in five Europe urban regions found a mean of 1.96 km^2. Similar results were found in a small study conducted with adolescents in Boston, where it was reported a mean area of 1.82 km^2. Stewart et al [46], found in a pilot study conducted in Auckland (New Zealand) an area of 3.54 km^2; Coulton et al [21], in a study with 6,224 adults in low-income communities in 10 US cities found an area of 2.33 km^2. A study conducted in Helsinki and Espoo, Finland, with 15,982 persons, that calculated the area inside the most visited points in a neighborhood, found an average area of 1.07 km^2 [47].

Despite the heterogeneities in the sampling and methods utilized between studies, researches that use maps found larger neighborhoods, indicating a possible relationship to the methodology used to access the perceived neighborhood. A possible explanation is that it is easier to remember important points in neighborhoods when participants look at a map. Using an open- or closed-ended question does not provide that kind of specific context.

The results of the multilevel model show us that there are contextual factors associated with perceived neighborhood domains. We found associations with the perceptions of the walking environment and with violence. The interpretation of the results of the scales should be done based on the analysis of the behavior of its domain [38]. The walking environment scale had highest values when the census tract had more people who reported that their neighborhoods have a physical environmental that encourage mobility and external activities. To our knowledge, the literature does not report a similar relationship, but it is plausible that an area that stimulates the mobility of people, facilitating diverse activities within the neighborhood could be also related with a large perception of neighborhood scale.

Violence scale, which reports higher values for more violent neighborhoods, was associated with larger perceived neighborhood scale. This appears contradictory at first glance, but people with larger perceived neighborhoods are likely to have greater social contact and exposure to the environment and may therefore be able to identify the problems within the neighborhood.

Regarding connectivity, we found an association between high street connectivity and larger perceived neighborhood scale only in a univariate analysis, despite a negative association found in another publication[21]. After adjustments, connectivity was no longer significant although the plausibility of the association; highly connected streets tend to be in busier places with a high demographic density and intense automobile traffic that hampers social contact and favors less extensive perceptions.

Demographic density was not associated with neighborhood perception.The literature consulted differs in relation to this variable. Some studies have found [21 23] association between smaller perceived neighborhood and greater population density. Others have reported an association between higher population densities and larger neighborhoods [24 48] and further one, such ours, found no relationship [26 45]. However, neighborhoods with a high population density, especially if car traffic is intense could also impoverished social contact among neighbors favoring a lower neighborhood perception, in the same direction of connectivity.

This study has specific limitations that need to be mentioned as listed: 1. The use of a closed-ended question to obtain the perceived neighborhood scale does not specify the spaces that individuals are actually exposed to; 2. The cross-sectional design of the study limits the interpretation of some results, due to the possibility of reverse causality; 3. The results of this

study are from a large urban center and are not necessarily valid for smaller cities and rural areas; 4. The findings may not apply to children since those younger than 18 years were not included in this study.

The identification of the contextual factors associated with the perception of neighborhood scale have important methodological implications, especially for studies that intend to investigate the association of social factors of the neighborhood with health events. The perceived neighborhood scale is a fundamental tool for the creation of more precise and coherent neighborhood boundaries informed by the places actually experienced by individuals.

One of the motivations of this study is related to the fact that a large amount of research in eco-epidemiology and community practice tends to use artificial definitions of neighborhoods boundaries. The results of this study demonstrate that there is heterogeneity among residents on their perceived neighborhood scale, reinforcing the argument that researchers need to use more personalized ways to define neighborhood boundaries. Most research uses census tracts as a proxy for neighborhoods due the availability of data aggregated to this level, but the increased use of GIS techniques supports more individualized neighborhood definitions that can be used to avoid problems regarding the choice of neighborhood size and its operationalization. A more carefully defined neighborhood unit will help future eco-epidemiological studies to produce evidence to support a community practices.

Contributors:

de Almeida Celio, Fabiano de Almeida: Fabiano participated in the design of the study, its conception. Fabiano was responsible for writing the article and the following activities: He setting up the database, he made the statistical analysis, he reviewed the text, and approved the final version.

Friche, Amélia Augusta de Lima: Professor Friche participated in the design and construction of the field work; reviewed all versions of thework, ensured its accuracy and integrity and she approved the final version.

Jennings, M. Zane: Zane participated in design; ensured the accuracy and integrity of the data; approved all versions of the work and the final version.

Andrade, Amanda Cristina de Souza: Amanda participated in the design of the work, review all versions, statistical analysis, ensured the accuracy and precision of the data, approved the final version.

Xavier, Cesar Coelho: Professor Xavier participated in the design and construction of the field work; reviewed all versions of the work, ensured its accuracy and integrity and he approved the final version.

Proietti, Fernando: Professor Proietti participated in the design and construction of the field work; reviewed all versions of the work, ensured itsaccuracy and integrity and he approved the final version.

Coulton, Claudia J.: Professor Coulton participated in design; ensured the accuracy and integrity of the data; approved all versions of the workand the final version.

Caiaffa, Waleska Teixeira: Professor Caiaffa participated in the design and construction of the field work; reviewed all versions of the work,ensured its accuracy and integrity and she approved the final version.

Competing Interests Declaration:
All authors declare they have no actual or potential competing financial interest.

Funding by:
National Health Fund of the Ministry of Health, Fapemig, CNPq, NIH / Fogarty International Center.

Data sharing:
All relevant data are within the paper. The data underlying this study are third party data and are available to all interested researchers. To submit a proposal the OSUBH Coordinating Center at http://site.medicina.ufmg.br/osubh/contato/.

References

1. Roux AVD. Health in cities: is a systems approach needed? Cadernos de Saúde Pública 2015;**31**:9-13
2. Cohen D, Spear S, Scribner R, et al. " Broken windows" and the risk of gonorrhea. American journal of public health 2000;**90**(2):230
3. Schaefer-McDaniel N, Caughy MOB, O'Campo P, et al. Examining methodological details of neighbourhood observations and the relationship to health: a literature review. Social science & medicine 2010;**70**(2):277-92
4. Diez Roux AV, Mair C. Neighborhoods and health. Annals of the New York Academy of Sciences 2010;**1186**(1):125-45
5. Roux AVD, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. New England Journal of Medicine 2001;**345**(2):99-106
6. Mair CF, Roux AVD, Galea S. Are neighborhood characteristics associated with depressive symptoms? A critical review. Journal of epidemiology and community health 2008:jech. 2007.066605
7. Andrade ACdS, Peixoto SV, Friche AAdL, et al. Social context of neighborhood and socioeconomic status on leisure-time physical activity in a Brazilian urban center: The BH Health Study. Cadernos de Saúde Pública 2015;**31**:136-47
8. Fernandes AP, Andrade ACdS, Ramos CGC, et al. Leisure-time physical activity in the vicinity of Academias da Cidade Program in Belo Horizonte, Minas Gerais State, Brazil: the impact of a health promotion program on the community. Cadernos de Saúde Pública 2015;**31**:195-207
9. Chaix B, Kestens Y, Bean K, et al. Cohort Profile: Residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases—The RECORD Cohort Study. International Journal of Epidemiology 2011;**107**:1-10
10. Ludwig J, Duncan GJ, Gennetian LA, et al. Neighborhood effects on the long-term well-being of low-income adults. Science 2012;**337**(6101):1505-10
11. Kaiser P, Diez Roux AV, Mujahid M, et al. Neighborhood environments and incident hypertension in the multi-ethnic study of atherosclerosis. American journal of epidemiology 2016;**183**(11):988-97
12. Johnson DA, Simonelli G, Moore K, et al. The neighborhood social environment and objective measures of sleep in the Multi-Ethnic Study of Atherosclerosis. Sleep 2017;**40**(1)
13. Chaix B, Merlo J, Evans D, et al. Neighbourhoods in eco-epidemiologic research: delimiting personal exposure areas. A response to Riva, Gauvin, Apparicio and Brodeur. Social science & medicine 2009;**69**(9):1306-10
14. Flowerdew R, Manley DJ, Sabel CE. Neighbourhood effects on health: does it matter where you draw the boundaries? Social science & medicine 2008;**66**(6):1241-55
15. Galster G. On the nature of neighbourhood. Urban studies 2001;**38**(12):2111-24
16. Santos SM, Chor D, Werneck GL. Demarcation of local neighborhoods to study relations between contextual factors and health. International journal of health geographics 2010;**9**(1):1
17. Clapp JM, Wang Y. Defining neighborhood boundaries: Are census tracts obsolete? Journal of Urban Economics 2006;**59**(2):259-84
18. Chaskin RJ. Perspectives on neighborhood and community: a review of the literature. The Social Service Review 1997:521-47
19. Chappell NL, Funk LM, Allan D. Defining community boundaries in health promotion

research. American Journal of Health Promotion 2006;**21**(2):119-26
20. Chaix B, Meline J, Duncan S, et al. GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? Health & Place 2013;**21**:46-51
21. Coulton C, Jennings MZ, Chan T. How Big is My Neighborhood? Individual and Contextual Effects on Perceptions of Neighborhood Scale. American Journal of Community Psychology 2012:1-11 doi: 10.1007/s10464-012-9550-6[published Online First: Epub Date]|.
22. Coulton CJ, Korbin J, Chan T, et al. Mapping residents' perceptions of neighborhood boundaries: a methodological note. American journal of community psychology 2001;**29**(2):371-83
23. Haney W, Knowles E. Perception of neighborhoods by city and suburban residents. Human Ecology 1978;**6**(2):201-14 doi: 10.1007/bf00889095[published Online First: Epub Date]|.
24. Charreire H, Feuillet T, Roda C, et al. Self□defined residential neighbourhoods: size variations and correlates across five European urban regions. Obesity Reviews 2016;**17**(S1):9-18
25. Guest AM, Lee BA. How urbanites define their neighborhoods. Population and Environment 1984;**7**(1):32-56
26. Pebley AR, Sastry N. Our place: perceived neighborhood size and names in Los Angeles. Los Angeles: California Center for Population Research, University of California 2009
27. Célio FdA, Xavier CC, Andrade ACdS, et al. Características individuais associadas à autopercepção da extensão territorial da vizinhança. Cadernos de Saúde Pública 2014;**30**:1935-46
28. Cutchin MP, Eschbach K, Mair CA, et al. The socio-spatial neighborhood estimation method: an approach to operationalizing the neighborhood concept. Health & place 2011;**17**(5):1113-21
29. Fotheringham AS, Wong DW. The modifiable areal unit problem in multivariate statistical analysis. Environment and planning A 1991;**23**(7):1025-44
30. Duncan DT, Kawachi I, Subramanian SV, et al. Examination of How Neighborhood Definition Influences Measurements of Youths' Access to Tobacco Retailers: A Methodological Note on Spatial Misclassification. American Journal of Epidemiology 2013 doi: 10.1093/aje/kwt251[published Online First: Epub Date]|.
31. Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing" neighborhood effects": Social processes and new directions in research. Annual review of sociology 2002:443-78
32. Hillier B. Spatial sustainability in cities: organic patterns and sustainable forms. 2009
33. Andrade RGd, Chaves OC, Costa DAdS, et al. Overweight in men and women among urban area residents: individual factors and socioeconomic context. Cadernos de Saúde Pública 2015;**31**:148-58
34. Camargos VP, César CC, Caiaffa WT, et al. Imputação múltipla e análise de casos completos em modelos de regressão logística: uma avaliação prática do impacto das perdas em covariáveis. Cad Saúde Pública 2011;**27**:2299-313
35. Ferreira AD, César CC, Malta DC, et al. Validade de estimativas obtidas por inquérito telefônico: comparação entre VIGITEL 2008 e Inquérito Saúde em Beagá. Rev Bras Epidemiol 2011;**14**(1):16-30
36. Gerência de Epidemiologia e Informação Secretaria Municipal de Saúde de Belo Horizonte. Índice de vulnerabilidade à saúde. Secondary Índice de vulnerabilidade à saúde 2003. http://www.pbh.gov.br/smsa/biblioteca/gabinete/risco2003.
37. Friche AA, Xavier CC, Proietti FA, et al. *Saúde urbana em Belo Horizote*. 1 ed, 2015.

38. Friche AA, Diez-Roux AV, Cesar CC, et al. Assessing the Psychometric and Ecometric Properties of Neighborhood Scales in. Journal of urban health : bulletin of the New York Academy of Medicine 2012 doi: 10.1007/s11524-012-9737-z[published Online First: Epub Date]|.
39. Song Y, Knaap G-J. Measuring urban form: Is Portland winning the war on sprawl? Journal of the American Planning Association 2004;**70**(2):210-25
40. Demográfico IC. Resultados do universo. Rio de Janeiro: IBGE 2010
41. Barros AJ, Victora CG. Indicador econômico para o Brasil baseado no censo demográfico de 2000. Revista de Saúde Pública 2005;**39**(4):523-29
42. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health 2006;**60**(4):290-7 doi: 10.1136/jech.2004.029454[published Online First: Epub Date]|.
43. Twisk J. Applied multilevel analysis: a practical guide. 2006: Cambridge University Press, Cambridge, UK.
44. StataCorp L. Stata survey data reference manual, 1985.
45. Sastry N, Pebley AR, Zonta M. Neighborhood definitions and the spatial dimension of daily life in Los Angeles. California Center for Population Research 2002
46. Stewart T, Duncan S, Chaix B, et al. A novel assessment of adolescent mobility: a pilot study. International journal of behavioral nutrition and physical activity 2015;**12**(1):18
47. Hasanzadeh K, Broberg A, Kyttä M. Where is my neighborhood? A dynamic individual-based definition of home ranges and implementation of multiple evaluation criteria. Applied Geography 2017;**84**:1-10
48. Vallée J, Le Roux G, Chaix B, et al. The 'constant size neighbourhood trap'in accessibility and health studies. Urban Studies 2014:0042098014528393

Table 1. Univariate analysis of individual variables by perceived neighborhood scale: percentages, means and standard deviations.

INDIVIDUAL VARIABLES	Perceived Neighborhood scale (1-4 and %)*				OR (CI 95 %) [1]	p
	1 (57.8%)	2 (23.3%)	3 (7.4%)	4 (11.5%)		
Gender (female)	56.4	51.0	45.9	45.9	0,70 (0,58 – 0,83)	<0.001
Employment state (working)	62.0	65.6	69	73.3	1.36 (1.14 - 1.61)	<0.001
Presence of child younger than 10 years (yes)	33.1	31.5	30.6	33.0	0.95 (0.82 - 1.11)	0.540
Number of relatives and friends living in the same neighborhood (almost all)	2.92	7.07	7.83	13.03	6.30 (4.00 - 9.92)	<0.001
Recognizes most of them people passing by the door of his/her house (yes)	8.0	12.6	13.1	21.3	5.55 (3.04 - 10.11)	<0.001
	Mean (Standard deviation)				OR (CI 95 %) [1]	p
Age (years)	44.7 (0.35)	44.9 (0.57)	41.1 (0.92)	43.2 (0.74)	0.99 (0.99 - 1.00)	0.060
Socioeconomic position (NEI) **	586.8 (4.06)	601.6 (6.6)	582.8 (11.4)	601.4 (0.0)	1.13 (1.04 - 1.23)	<0.001
Time of residence in the same neighborhood (years)	14.8 (0.26)	16.8 (0.44)	16.2 (0.71)	16.6 (0.60)	1.01 (1.01 -1.02)	<0.001

* (1) up to the block or street you live on; (2) within 5 blocks; (3) within ten blocks; and (4) more than ten blocks away; OR – odds ratio; CI 95% – 95% confidence interval; NEI - national economic index; ** Odds ratio calculated based on an increase of 200 NEI points; [1] estimated by a multilevel ordinal logistic regression model, the reference category is the smaller neighborhood.

Table 2. Univariate analysis of contextual variables by perceived neighborhood scale: percentages, means and standard deviations.

CONTEXTUAL VARIABLES	NEIGHBORHOOD EXTENSION SCALE (1-4 and %)*				OR (CI 95 %) [1]	p
	1 (57.8%)	2 (23.3%)	3 (7.4%)	4 (11.5%)		
Connectivity [2]						
0 to 3	39.91	40.91	41.87	40.37	1.00	
4	23.92	23.42	36.11	25.28	1.04 (0.83 - 1.30)	0.760
5 to 9	37.17	35.63	22.02	34.34	0.85 (0.70 - 1.04)	0.120
	Mean (Standard deviation)				OR (CI 95 %) [1]	p
Aesthetic Quality scale	2.96 (0.03)	3.07 (0.04)	3.06 (0.07)	3.09 (0.04)	1.23 (1.03 - 1.46)	0.020
Walking Environment scale	3.20 (0.01)	3.24 (0.02)	3.27 (0.02)	3.28 (0.02)	3.37 (2.09 - 5.44)	<0.001
Violence scale	1.90 (0.02)	1.95 (0.03)	1.89 (0.05)	2.00 (0.04)	1.20 (1.03 - 1.40)	0.020
Safety scale	2.96 (0.03)	2.93 (0.05)	2.89 (0.09)	2.86 (0.05)	0.92 (0.79 - 1.08)	0.190
Population density (per square kilometer)	12487.35 (791.26)	11740.34 (704.83)	12274.31 (740.63)	12627.22 (865.05)	1.00 (0.99 - 1.01)	0.650

* (1) up to the block or street you live on; (2) within 5 blocks; (3) within ten blocks; and (4) more than ten blocks away; OR – odds ratio; CI 95% – 95% confidence interval. [1] estimated by a multilevel ordinal logistic regression model, the reference category is the smaller neighborhood; [2] 0 means poorly connected streets and 9 heavily connected streets.

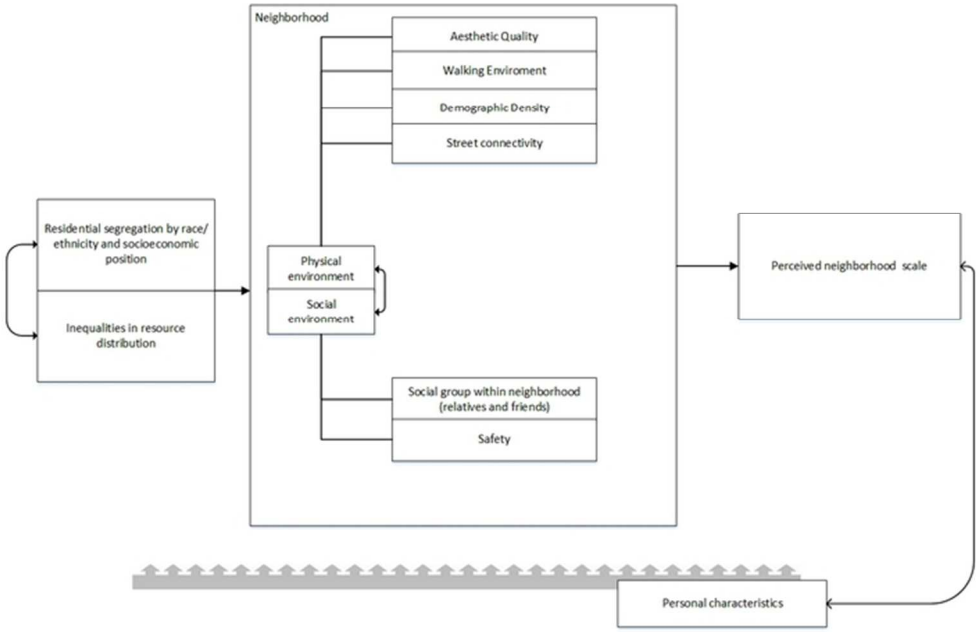
Table 3. Multilevel ordinal logistic regression for the resident perceived neighborhood scale

Variables	Null Model	Contextual variables		Contextual variables + individual variables	
		OR (CI 95%) [1]	p	OR (CI 95%) [1]	p
Aesthetic Quality scale		1.21 (0.97 - 1.41)	0.060	1.13 (0.92 -1.39)	0.230
Walking Environment scale		2.96 (1.71 - 5.13)	<0.001	2.22 (1.29 -3.82)	<0.001
Violence scale		1.35 (1.12 - 1.62)	<0.001	1.23 (1.01 - 1.51)	0.040
Safety scale		0.97 (0.82 - 1.14)	0.710	0.99 (0.83 -1.19)	0.950
Connectivity [2]					
4		1.00 (0.80 -1.26)	0.940	1.06 (0.85 - 1.34)	0.570
5 e 9		0.82 (0.67 - 1.01)	0.060	0.89 (0.72 - 1.11)	0.310
INDIVIDUAL					
Age (years)				1.00 (0.99 - 1.00)	0.350
Gender (female)				0.81 (0.68 - 0.96)	0.020
Number of relatives and friends living in the same neighborhood (almost all)				4.63 (2.84 - 7.57)	<0.001
Recognizes most of them people passing by the door of his/her house (yes)				3.33 (1.72 - 6.25)	<0.001
Employment state (working)				1.26 (1.06 - 1.50)	0.010
Time of residence in the same neighborhood (years)				1.01 (1.00 - 1.01)	0.130
Presence of child younger than 10 years old (yes)				0.97 (0.81 - 1.16)	0.740
Socioeconomic position**				1.17 (1.06 - 1.29)	<0.001
MODEL INFORMATION					
Variance	0.2567		0.2292		0.2336
MOR	1.62		1.58		1.58
Proportional change in variance	-		10.71		9.00
AIC	8749.26		8668.44		8091.83

OR – odds ratio; CI 95% – 95% confidence interval; NEI - national economic index; MOR - median values of the odds ratio; AIC - Akaike Information Criterion; ** Odds ratio calculated based on an increase of 200 NEI points; [1] estimated by a multilevel ordinal logistic regression model, the reference category is the smaller neighborhood; [2] 0 means connected streets and 9 heavily connected streets.

Figure Legend:

Figure 1- Theoretical model for factors associated with perceived neighborhood scale



Theoretical Model for perceived neighborhood scale

30x19mm (600 x 600 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation	
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	x
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	x
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	x
Objectives	3	State specific objectives, including any prespecified hypotheses	x
Methods			
Study design	4	Present key elements of study design early in the paper	x
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	x
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	x
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	x
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Not applicable.
Bias	9	Describe any efforts to address potential sources of bias	Not applicable.
Study size	10	Explain how the study size was arrived at	x
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	x
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	X
		(*b*) Describe any methods used to examine subgroups and interactions	Not applicable.
		(*c*) Explain how missing data were addressed	Not applicable.
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	x
		(*e*) Describe any sensitivity analyses	Not applicable.
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	x
		(b) Give reasons for non-participation at each stage	Not applicable.
		(c) Consider use of a flow diagram	Not applicable.
Descriptive data	14*	(a) Give characteristics of study participants (eg	x

		demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	Not applicable.
Outcome data	15*	Report numbers of outcome events or summary measures	x
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	x
		(*b*) Report category boundaries when continuous variables were categorized	x
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Not applicable.
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Not applicable.
Discussion			
Key results	18	Summarise key results with reference to study objectives	x
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	x
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	x
Generalisability	21	Discuss the generalisability (external validity) of the study results	x
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	x

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Contextual characteristics associated with the perceived neighborhood scale, in a cross-sectional study in a large urban center in Brazil

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-021445.R2
Article Type:	Research
Date Submitted by the Author:	13-Jun-2018
Complete List of Authors:	de Almeida Celio, Fabiano; Universidade Federal de Minas Gerais, Public health Friche, Amélia Augusta ; Universidade Federal de Minas Gerais Faculdade de Medicina, Public health Jennings, M. Zane; Case Western Reserve University Jack Joseph and Morton Mandel School of Applied Social Sciences Andrade, Amanda Cristina; Universidade Federal de Minas Gerais Faculdade de Medicina, public health Xavier, Cesar ; faculdade de saúde e ecologia humana Proietti, Fernando; Faculdade de Saúde e Ecologia Humana; Centro de Pesquisas Rene Rachou Coulton, Claudia J.; Case Western Reserve University Jack Joseph and Morton Mandel School of Applied Social Sciences Caiaffa, Waleska; Federal University of Minas Gerais, Medicine
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Public health
Keywords:	neighborhood, urban health, Perceived Neighborhood, neighborhood scale, Community perceptions

Contextual characteristics associated with the perceived neighborhood scale, in a cross-sectional study in a large urban center in Brazil

Corresponding author: Fabiano de Almeida Célio, maitiz@gmail.com**, +55 31 99877 7761**

1) **Fabiano de Almeida Célio**[1], maitiz@gmail.com; 2) **Amélia Augusta de Lima Friche**[1], gutafriche@gmail.com; 3) **M. Zane Jennings**[2], mzj@case.edu; 4) **Amanda Cristina de Souza Andrade**[1], amandasouza_est@yahoo.com.br; 5) **Cesar Coelho Xavier**[3], cesarcxavier@gmail.com; 6) **Fernando Augusto Proietti**[3], fernandoaproietti@gmail.com; 7) **Claudia J. Coulton**[2], claudia.coulton@case.edu; 8) **Waleska Teixeira Caiaffa**[1], caiaffa.waleska@gmail.com

1. Observatory for Urban Health in Belo Horizonte, School of Medicine, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil
2.Case Western Reserve University Jack Joseph and Morton Mandel School of Applied Social Sciences, Cleveland OH, USA
3. Faculdade de Saúde e Ecologia Humana, Vespasiano, Minas Gerais, Brazil
4. Centro de Pesquisas René Rachou, Belo Horizonte, Minas Gerais, Brazil

Keywords: neighbourhood, urban health; perceived neighbourhood; neighbourhood scale, community perceptions

Word count: 3814

Reference count: 47

Abstract

Introduction: Health outcomes have been associated with physical and social characteristics of neighbourhoods, but little is known about the relationship between contextual factors and perceived neighbourhood scale. **Objective:** To identify the contextual factors associated with self-perceived neighbourhood scale. **Methods:** We analysed data from a cross-sectional population-based study in Belo Horizonte, Brazil, that took place in 2008-2009. The dependent variable was perceived neighbourhood, encoded as an ordinal scale based upon a brief description of the concept of the neighbourhood, and two independent scales relating distance, expressed in terms of geography and time. Street connectivity, demographic density and residents' perceptions of the neighbourhoods' physical and social environment were used as contextual predictors. Individual characteristics were used as covariates. Multilevel ordinal logistic regression models estimated the association between perceived neighbourhood scale and contextual characteristics. **Results:** Residents that perceive better walkability (OR=2.96; CI 95%: 1.29 - 3.82) and high amounts of violence (OR=1.35; CI 95%: 1.12 - 1.62) perceived their neighbourhoods to be larger, even after adjusting for individual characteristics.

Conclusion: There are contextual factors that are associated with self-perceived neighbourhood scale. Careful definition of neighbourhood scale is a key factor in improving the results of eco-epidemiological studies. Although these findings must be further explored in other studies, these results can contribute to a better understanding of an appropriate choice of neighbourhood scale, especially for cities in Latin America.

Summary Boxes:

Strengths and limitations of this study:

- Large sample comes from an urban centre in Latin America
- Analysis includes individual and contextual factors
- Neighbourhood definition can be obtained by closed-ended questions
- Analysis could identify contextual factors associated with perceived neighbourhood scale
- Analysis takes into account physical and social factors of the neighbourhood

Introduction

Eco-epidemiological research has increasingly used the concept of neighbourhood as the geographical area within which physical and social environmental features affect individuals' health outcomes, as part of an emphasis on a more holistic understanding of the factors and processes shaping health outcomes within urban areas [1].

Features in the neighbourhood help explain inequalities in health, can be used in studies aiming to evaluate community interventions intended to improve health outcomes [2-4], and have been shown to be predictive of health outcomes and health-affecting behaviours, such as cardiovascular diseases [5], sexually transmitted diseases [2], mental illness [6], and physical activity [7 8], among others [9-12].

However, the neighbourhood is a complex concept, and its definitions in epidemiological studies vary widely [13] and have different methodological approaches [14]. Chaix et al (2009) describe two approaches for defining neighbourhood in epidemiological research: the territorial neighbourhood and the ego-centred neighbourhood approaches.

Territorial neighbourhoods are generally administrative areas corresponding to a territory-subdividing approach. However, more complex definitions of territorial neighbourhoods may consider built environment features and population characteristics. Researchers using this approach often select administratively defined, mutually exclusive geographic units, such as census tracts or municipal boundaries, as proxies for neighbourhoods [13 15]. Assuming resident homogeneity [16 17], this approach is adopted because secondary data is often easily available and spatial references are obtainable, which facilitates reproducibility and comparability across studies or over time. However, territorial neighbourhoods consider the same areas for different individuals, and thus, individual differences in neighbourhood experience and exposure cannot be captured under this approach [16 17]. When the same area is attributed to several individuals in a given area, the potential for error is introduced because individuals may not be exposed in a homogeneous way to the physical and social environment of the territory.

The second approach is called ego-centred neighbourhoods and is based upon the idea that the contextual factors affecting individuals will differ depending on the actual location and particular geographic circumstances of those individuals. Several techniques can be used to define this approach. Most importantly, the ego-centred neighbourhood results in neighbourhoods that may overlap, are not mutually exclusive, and are specific to the

household or individual resident [13]. This approach can be operationalized in three different ways. One uses a buffer, generally a circular area centred upon the individual's residence, resulting in neighbourhoods of the same size, though made up of different areas, that may overlap with one another but are not identical. The second approach involves using individual behavioural activity spaces measured by GPS. This approach captures each individual's movements and activities, creating a unique measure of contextual exposure [9 18]. The third method relies on individuals' perceived neighbourhoods.

Perceived neighbourhoods, in turn, can be identified by different strategies. Residents may be asked to identify or draw their neighbourhood on a map [19-22], or, alternatively, researchers may ask residents how large they consider their neighbourhood to be or how long it takes to walk from the resident's house to the end of their neighbourhood [13 23-25]. This last technique has the advantage of being easily understood by residents and quickly and inexpensively conducted by researchers.

Regardless of the methods, neighbourhood scale needs to be carefully considered. When it is not correctly operationalized and defined, the measures derived can be considered problematic and questionable. Consequently, the understanding of health impacts through the lens of the neighbourhood can be undermined [26]. One problem that may arise is known in geography as the [27] "modifiable area unit problem". Aggregating epidemiological data into differently sized territorial units can yield varying exposure measure results, making it difficult or even impossible to compare findings. Generally, the error of choice of territorial unit of analysis is non-differential, which may underestimate association measures or even not find associations when they do exist [28].

The attributes that make the neighbourhood of an individual a singular place are commonly characterized by the following qualities: a) social interaction; b) social norms and collective effectiveness; c) institutional resources (schools, health facilities and others); and d) routine activities within the neighbourhood. As we can see, it is difficult not to incur some kind of neighbourhood boundary definition error when the internal dynamics of the place under study are unknown [29].

Perceived neighbourhood scale has been found to be related to individual characteristics, such as socioeconomic position, employment, evaluation of the aesthetic aspects, number of relatives living in the same neighbourhood and familiarity with many people in the neighbourhood [25]. However, the scale of perceived neighbourhood can be influenced by

contextual factors such as population density, land use patterns, and collective efficacy [19]. The connectivity of the streets that directly influence the number of routes available to the various points of interest within a neighbourhood can also influence the perception of its size, because connectivity may change the way residents use and circulate in physical space [30].

This work, by investigating perceived neighbourhood scale, addresses an important methodological question, which concerns the appropriate scale of territorial units of analysis, reducing possible errors inherent to the process of investigating neighbourhood impact on health outcomes. Despite research results indicating a relationship between neighbourhood and health, it is still rare to find studies that measure the influence of contextual factors as shaped by perceived neighbourhood scale. In Latin America, we have not found any studies with this same purpose. Therefore, the objective of this study is to analyse the context attributes associated with the perceived neighbourhood scale in a large urban centre in Brazil.

Methods

Data and Sample

The data for this study come from a cross-sectional population-based study called *BH Health Study*, conducted by the Belo Horizonte Observatory for Urban Health (OSUBH) in 2008-2009 and nested in the Federal University of Minas Gerais (UFMG). The participants of the study were residents belonging to two of the nine sanitary districts of Belo Horizonte: Barreiro and West. These districts were selected because they presented heterogeneity within the city in relation to social, sociodemographic and health indicators [31-33].

A stratified sample was selected in a three-stage process. To ensure the representation of residents of all socioeconomic levels, the study area was stratified by the health vulnerability index [34], a geocoded index created by combining social, demographic, economic, and health indicators from different sources for each census tract. At the end of the first and second steps of the sampling process, 149 census tracts and 4,048 households were randomly selected. In the third stage, one resident over 18 years old was randomly selected in each of the identified households [35].

Contextual predictors

The dependent variable for this study is the perceived neighbourhood scale, which was originally encoded as an ordinal variable with 7 options. To obtain the scale, the interviewer read a brief description of the concept of the neighbourhood: "The neighbourhood is the place

where you live and perform routine tasks, such as going to the bakery, grocery store, and local businesses; visiting your neighbours; and walking. The neighbourhood can be understood as the area where you recognize most of the people". Then, the interviewee was asked, "Thinking of your neighbourhood, would you describe it as including the following: (1) the houses next door? (2) the block or street you live on? (3) the area within 5 blocks? (4) the area within ten blocks? (5) the area more than ten blocks away? (6) your neighbourhood? (7) your neighbourhood and nearby neighbourhoods?" Subsequently, this variable was recoded, using as reference an additional measure of neighbourhood scale. This measure was a continuous variable obtained from the following question: "How much time in minutes would you spend walking from the door of your house to the end of what you consider your neighbourhood?" The mean walking time obtained for each of the seven options of the first ordinal question variable was used to collapse the final dependent variable into four options. This procedure was adopted by considering the non-overlapping portion of the 95% confidence intervals (95% CI) between each stratum. Thus, the outcome variable called the perceived neighbourhood scale was recoded into four categories: (1) up to the block or street you live on; (2) within 5 blocks; (3) within ten blocks; and (4) more than ten blocks away.

Independent variables

The independent variables were chosen based on the theoretical (Figure 1) model using other studies [4 25]. The variables relating to the physical and social environment of the neighbourhood were obtained from domains created by Friche et al [36]. Aggregated for each census tract, the domains provide a continuous score ranging from 1 to 4. In this study, we used the following domains: aesthetic quality, walking environment, safety and violence.

The aesthetic quality domain was obtained by asking the participants the following questions about their neighbourhood: 1) Is there trash or litter on the streets and sidewalks? 2) Is pleasant for children? 3) Is pleasant for young children and adolescents? 4) Are there trees that make the environment pleasant?

The walking environment domain was obtained by asking the participants the following about their neighbourhood: 1) How do you evaluate public places for sports and leisure? 2) How do you evaluate the traffic? 3) Are there stores at a distance you can walk? 4) Is it easy to walk? 5) How often do you see other people walking? 6. How often do you see other people exercising? 7) Do you feel safe walking during the day?

The violence domain was composed of the following questions: During the past 12 months,

did you see or hear about the following: 1) people being mugged in the neighbourhood streets? 2) people fighting using weapons? 3) people being killed by guns? 4) people being victims of sexual violence? 5) women of the neighbourhood being beaten by their husbands and/or partners or relatives? and 6) children or adolescents of the neighbourhood being assaulted or victims of violence perpetrated by their parents?

The safety domain was built with the following questions: In your neighbourhood, 1) you feel safe walking during the night; 2) violence is a problem.

This study also used contextual variables from census tracts and those collected by the city hall for administrative purposes. Street connectivity drawn from all street segments of the area in the study was obtained using Dephmap [37] (Space Syntax Ltd. University of London) software. This software handles the street segment as if it were an axial line and quantifies the segments that intersect each of these lines [30]. The software delivers a score between 0 and 9, where 0 represents streets with low connectivity and 9 represents highly connected streets [38]. The final variable was skewed, with a low prevalence of extreme values, so it was recoded into three categories: low connectivity (0 to 3); medium connectivity (4), and high connectivity (5 to 9).

Population density was calculated for each census tract using data from the 2010 National Census [39].

Individual variables

Individual characteristics were included as covariates that had been found to be predictors of neighbourhood scale in previous studies [19 25]. These characteristics included the following: gender; age (in years); employment status; length of residence in home (in years); presence of children under 10 years of age in the household; number of relatives in the same neighbourhood (none to all); number of people who pass in front of participants' houses who are known to them (none to all); and a composite indicator named the national economic index (NEI), which depicts the current socioeconomic position of the individual [40], based upon consumer goods instead of income.

Statistical analyses

A descriptive analysis was carried out, followed by an analysis of the association between size of the neighbourhood scale and contextual features estimated by a multilevel ordinal logistic regression model. The first level consisted of the individual-level variables, and the

second level consisted of the neighbourhood-level variables.

A regression model with random intercepts with a logit function were used to estimate the odds ratio (OR) and the confidence interval (CI 95%) [41]. The median values of the odds ratio (MOR) and the percentage of variance reduction were calculated. The Akaike information criterion (AIC) was used to compare models, with the model with the lowest AIC selected as the best model [42].

First, a null model (only the random intercept) was estimated to assess the contextual effect, and then a univariate analysis was performed with a multilevel ordinal logistic regression for each of the contextual variables. Second, independent domains with a coefficient that was significant at $p \leq 0.20$ (aesthetic quality, walking environment, violence domain and safety) in the univariate analysis were included as level 2 variables in the multiple analysis. Finally, we added the individual characteristics (age, gender, employment status, number of parents and friends in the neighbourhood, recognition of people passing by the door of your house, length of residence in the same neighbourhood, presence of children younger than 10 and socioeconomic position) at level 1 for adjustment.

The analyses were performed in the software STATA (Stata Corp., College Station, Texas), version 12.0. For all analyses, we used the svy command [43], which considers complex design and sampling weights. For all models, we considered a significance level of 5%.

Patient and Public Involvement

The research participants were selected in two regions of the city that show great internal heterogeneity in relation to sociodemographic characteristics. The local population was previously informed about the objectives and importance of the research through several approaches, including the involvement of community leader representatives, religious groups, school educators, and health agents of family and community health programmes. After the selection of the households, the objectives of the research were presented to each participant. The results of the study were thoroughly disseminated within the population and discussed with public policy administrators of the municipality.

Ethical issues

The study was approved by the Research Ethics Committee of the UFMG through opinion ETIC n° 253/06. All participants provided informed consent.

Results

The final sample had 4,048 respondents, 53.1% of whom were men and 46.9% of whom were women, with ages varying between 18 and 95 years (mean = 44.4 , SD = 16.9). We found that 57.8% of the participants considered their neighbourhood to extend from their own house to the end of the block, 23.3% considered their neighbourhood to be within the 5 closest blocks, 7.4% considered their neighbourhood to be within the nearest 10 blocks, and 11.5% considered their neighbourhood to be larger than 10 blocks from their home (table 1).

There was a linear relationship between the size of the perceived neighbourhood and the time to walk to his/her end of the neighbourhood, with the following average times, in minutes, for each neighbourhood size stratum: 6.1, 13.5, 19.8 and 29.2.

The distribution of street connectivity was almost the same, with 39.7% of streets with connectivity between 0 and 3; 24.8% with a value of 4; and 35.4% with values between 5 and 9. The mean population density was 12,264 residents/km^2 (685.9). All of these results are shown in table 2.

The multilevel model analysis began with the null model. The perceived neighbourhood size had significant variation within the census tract, based on the likelihood-ratio test ($p < 0.001$). The analysis showed that the following domains and variables were significantly associated with self-perceived neighbourhood size (table 3): walking environment (OR=2.96; CI 95%: 1.29 – 3.82), violence (OR=1.35; CI95%: 1.12 – 1.62), female gender (OR= 0.81; CI 95%: 0.68 – 0.96), greater number of relatives living in the neighbourhood (OR= 4.63; CI 95%: 2.84 – 7.57), recognition of more people in the neighbourhood (OR= 3.33; CI 95%: 1.72 – 6.25), and socioeconomic position (NEI) (OR= 1.17; CI 95%: 1.06 – 1.29).

Based on the Akaike information criterion (AIC), the best models were adjusted by individual variables at level 1. The median value of the odds ratio comes from the median value between the area with the highest odds of a larger perceived neighbourhood scale and the area with the lowest odds when randomly selecting two areas [41]. We found an MOR of 1.62 for the null model, 1.58 for the model with the contextual predictors, and 1.58 for the model adjusted by the individual variables. The results of the proportional change in variance show that the contextual predictors explained 10.7% of the total variance, and the model with contextual and individual variables explained 9.0%. Although the best model based on the AIC is the model with contextual and individual variables, the model with only contextual variables has more variation than the null model. This finding indicates that 10.0% of the contextual variance of perceived neighbourhood scale was attributed to the contextual factors and that

when we added the individual-level variables, it decreased slightly to 9.0% (table 3).

Discussion

Perceived neighbourhood scale was relatively small for many residents: 57.8% of the participants considered their neighbourhood to be residences closest to their home until the end of the block. Additionally, contextual factors, such as perceived quality of environmental conditions for walking and indicators of a violent environment, were associated with a larger perceived neighbourhood scale, even adjusted by individual-level variables.

These findings allow for comparison with those of previous studies, but care must be taken because each study has different approaches in measuring perceived neighbourhood. A study carried out in the city of Los Angeles [44] reported that 35.1% of the interviewees considered their neighbourhood the block or street that they live on, 25.0% several blocks or streets in each direction, 28.8% an area within a 15-minute walk, and 13.1% an area larger than a 15-minute walk. In other words, 86.9% of participants considered their neighbourhood an area larger than a 15-minute walk. This result is very similar to what we found; when we look at the average time taken to leave the self-perceived neighbourhood within each stratum, 81.1% of participants considered their neighbourhood an area smaller than a 15-minute walk.

Another study conducted in different areas of Seattle [23] found that 46.4% of participants considered their neighbourhood to extend from their own residential unit to no more than one block in each direction.

However, studies using maps as an approach to measure the perceived neighbourhood found much larger neighbourhoods. A study [22] conducted in five European urban regions found a mean perceived neighbourhood of 1.96 km^2. Similar results were found in a small study conducted with adolescents in Boston, where a mean area of 1.82 km^2 was reported. In a pilot study conducted in Auckland (New Zealand), Stewart et al [45] found a perceived neighbourhood area of 3.54 km^2; in a study with 6,224 adults in low-income communities in 10 US cities, Coulton et al [19] found an area of 2.33 km^2. A study conducted with 15,982 persons, in Helsinki and Espoo, Finland, that calculated the area inside the most visited points in a neighbourhood, found an average area of 1.07 km^2 [46].

Despite the heterogeneities in the sampling and methods utilized among studies, studies that used maps found larger neighbourhoods, indicating a possible relationship with the methodology used to access the perceived neighbourhood. A possible explanation is that it is

easier to remember important points in neighbourhoods when participants look at a map. Using an open- or closed-ended question does not provide that kind of specific context.

The results of the multilevel model show us that there are contextual factors associated with perceived neighbourhood domains. We found associations with the perceptions of the walking environment and with violence. The interpretation of the results of the scales should be performed based on the analysis of the behaviour of its domain [36]. The walking environment scale had highest values when the census tract had more people who reported that their neighbourhoods have a physical environment that encourages mobility and external activities. To our knowledge, the literature does not report a similar relationship, but it is plausible that an area that stimulates the mobility of people, facilitating diverse activities within the neighbourhood, could also be related to a large perceived neighbourhood scale.

The violence domain, which reports higher values for more violent neighbourhoods, was associated with larger perceived neighbourhood scale. This finding appears to be contradictory at first glance, but people with larger perceived neighbourhoods are likely to have greater social contact and exposure to the environment and may therefore be able to identify the problems within the neighbourhood.

Regarding connectivity, we found an association between high street connectivity and larger perceived neighbourhood scale only in a univariate analysis, despite a negative association found in another publication[19]. After adjustments, connectivity was no longer significant, although the plausibility of the association remains; highly connected streets tend to be located in busier places with a high demographic density and intense automobile traffic, which hampers social contact and favours less extensive perceptions.

Demographic density was not associated with neighbourhood perception. The literature consulted differs in relation to this variable. Some studies have found [19 21] an association between smaller perceived neighbourhood and greater population density. Others have reported an association between higher population densities and larger neighbourhoods [22 47], and yet other studies, such ours, have found no relationship [24 44]. However, neighbourhoods with a high population density, especially if car traffic is intense, could also have impoverished social contact among neighbours, favouring a lower neighbourhood perception, in the same direction of connectivity.

This study has specific limitations that need to be mentioned. First, the use of a closed-ended question to obtain the perceived neighbourhood scale does not specify the spaces to which

individuals are actually exposed. Second, the cross-sectional design of the study limits the interpretation of some results due to the possibility of reverse causality. Third, the results of this study are from a large urban centre and are not necessarily valid for smaller cities and rural areas. Fourth, the findings may not apply to children, since individuals younger than 18 years were not included in this study.

The identification of the contextual factors associated with the perception of neighbourhood scale have important methodological implications, especially for studies that intend to investigate the association between social factors of the neighbourhood and health events. The perceived neighbourhood scale is a fundamental tool for the creation of more precise and coherent neighbourhood boundaries informed by the places actually experienced by individuals.

One of the motivations of this study is related to the fact that a large amount of research in eco-epidemiology and community practice tends to use artificial definitions of neighbourhoods' boundaries. The results of this study demonstrate that there is heterogeneity among residents on their perceived neighbourhood scale, reinforcing the argument that researchers need to use more personalized ways to define neighbourhood boundaries. Most research uses census tracts as a proxy for neighbourhoods due to the availability of data aggregated at this level, but the increased use of GIS techniques supports more individualized neighbourhood definitions that can be used to avoid problems regarding the choice of neighbourhood size and its operationalization. A more carefully defined neighbourhood unit will help future eco-epidemiological studies to produce evidence to support community practices.

Contributors:

de Almeida Celio, Fabiano de Almeida: Fabiano participated in the design of the study and its conception. Fabiano was responsible for writing the article and the following activities: setting up the database, performing the statistical analysis, reviewing the text and approving the final version.

Friche, Amélia Augusta de Lima: Professor Friche participated in the design of the field work, reviewed all versions of the paper, ensured its accuracy and integrity and approved the final version.

Jennings, M. Zane: Zane participated in design, ensured the accuracy and integrity of the data, and approved the versions of the paper and the final version.

Andrade, Amanda Cristina de Souza: Amanda participated in the design of the work, reviewed all versions, performed the statistical analysis, ensured the accuracy and precision of the data, and approved the final version of the paper.

Xavier, Cesar Coelho: Professor Xavier participated in the design of the field work and approved the final version of the paper.

Proietti, Fernando: Professor Proietti participated in the design of the field work and approved the final version of the paper.

Coulton, Claudia J: Professor Coulton participated in the analysis of the paper and approved the final version.

Caiaffa, Waleska Teixeira: Professor Caiaffa participated in the design and construction of the field work, reviewed all versions of the paper, ensured its accuracy and integrity and approved the final version.

Competing Interests Declaration:
All authors declare that they have no actual or potential competing financial interest.

Funding by:
National Health Fund of the Ministry of Health, Fapemig, CNPq, NIH/Fogarty International Centre.

Data sharing:
All relevant data are within the paper. The data underlying this study are third-party data and are available to all interested researchers. To gain access to these data, please submit a proposal the OSUBH Coordinating Centre at http://site.medicina.ufmg.br/osubh/contato/.

References

1. Roux AVD. Health in cities: is a systems approach needed? Cadernos de Saúde Pública 2015;**31**:9-13
2. Cohen D, Spear S, Scribner R, et al. " Broken windows" and the risk of gonorrhea. American journal of public health 2000;**90**(2):230
3. Schaefer-McDaniel N, Caughy MOB, O'Campo P, et al. Examining methodological details of neighbourhood observations and the relationship to health: a literature review. Social science & medicine 2010;**70**(2):277-92
4. Diez Roux AV, Mair C. Neighborhoods and health. Annals of the New York Academy of Sciences 2010;**1186**(1):125-45
5. Roux AVD, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. New England Journal of Medicine 2001;**345**(2):99-106
6. Mair CF, Roux AVD, Galea S. Are neighborhood characteristics associated with depressive symptoms? A critical review. Journal of epidemiology and community health 2008:jech. 2007.066605
7. Andrade ACdS, Peixoto SV, Friche AAdL, et al. Social context of neighborhood and socioeconomic status on leisure-time physical activity in a Brazilian urban center: The BH Health Study. Cadernos de Saúde Pública 2015;**31**:136-47
8. Fernandes AP, Andrade ACdS, Ramos CGC, et al. Leisure-time physical activity in the vicinity of Academias da Cidade Program in Belo Horizonte, Minas Gerais State, Brazil: the impact of a health promotion program on the community. Cadernos de Saúde Pública 2015;**31**:195-207
9. Chaix B, Kestens Y, Bean K, et al. Cohort Profile: Residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases—The RECORD Cohort Study. International Journal of Epidemiology 2011;**107**:1-10
10. Ludwig J, Duncan GJ, Gennetian LA, et al. Neighborhood effects on the long-term well-being of low-income adults. Science 2012;**337**(6101):1505-10
11. Kaiser P, Diez Roux AV, Mujahid M, et al. Neighborhood environments and incident hypertension in the multi-ethnic study of atherosclerosis. American journal of epidemiology 2016;**183**(11):988-97
12. Johnson DA, Simonelli G, Moore K, et al. The neighborhood social environment and objective measures of sleep in the Multi-Ethnic Study of Atherosclerosis. Sleep 2017;**40**(1)
13. Chaix B, Merlo J, Evans D, et al. Neighbourhoods in eco-epidemiologic research: delimiting personal exposure areas. A response to Riva, Gauvin, Apparicio and Brodeur. Social science & medicine 2009;**69**(9):1306-10
14. Flowerdew R, Manley DJ, Sabel CE. Neighbourhood effects on health: does it matter where you draw the boundaries? Social science & medicine 2008;**66**(6):1241-55
15. Galster G. On the nature of neighbourhood. Urban studies 2001;**38**(12):2111-24
16. Chaskin RJ. Perspectives on neighborhood and community: a review of the literature. The Social Service Review 1997:521-47
17. Chappell NL, Funk LM, Allan D. Defining community boundaries in health promotion research. American Journal of Health Promotion 2006;**21**(2):119-26
18. Chaix B, Meline J, Duncan S, et al. GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? Health & Place 2013;**21**:46-51
19. Coulton C, Jennings MZ, Chan T. How Big is My Neighborhood? Individual and Contextual Effects on Perceptions of Neighborhood Scale. American Journal of Community Psychology 2012:1-11 doi: 10.1007/s10464-012-9550-6[published Online First: Epub Date]|.
20. Coulton CJ, Korbin J, Chan T, et al. Mapping residents' perceptions of neighborhood boundaries: a methodological note. American journal of community psychology 2001;**29**(2):371-83

21. Haney W, Knowles E. Perception of neighborhoods by city and suburban residents. Human Ecology 1978;**6**(2):201-14 doi: 10.1007/bf00889095[published Online First: Epub Date]|.
22. Charreire H, Feuillet T, Roda C, et al. Self-defined residential neighbourhoods: size variations and correlates across five European urban regions. Obesity Reviews 2016;**17**(S1):9-18
23. Guest AM, Lee BA. How urbanites define their neighborhoods. Population and Environment 1984;**7**(1):32-56
24. Pebley AR, Sastry N. Our place: perceived neighborhood size and names in Los Angeles. Los Angeles: California Center for Population Research, University of California 2009
25. Célio FdA, Xavier CC, Andrade ACdS, et al. Características individuais associadas à autopercepção da extensão territorial da vizinhança. Cadernos de Saúde Pública 2014;**30**:1935-46
26. Cutchin MP, Eschbach K, Mair CA, et al. The socio-spatial neighborhood estimation method: an approach to operationalizing the neighborhood concept. Health & place 2011;**17**(5):1113-21
27. Fotheringham AS, Wong DW. The modifiable areal unit problem in multivariate statistical analysis. Environment and planning A 1991;**23**(7):1025-44
28. Duncan DT, Kawachi I, Subramanian SV, et al. Examination of How Neighborhood Definition Influences Measurements of Youths' Access to Tobacco Retailers: A Methodological Note on Spatial Misclassification. American Journal of Epidemiology 2013 doi: 10.1093/aje/kwt251[published Online First: Epub Date]|.
29. Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing" neighborhood effects": Social processes and new directions in research. Annual review of sociology 2002:443-78
30. Hillier B. Spatial sustainability in cities: organic patterns and sustainable forms. 2009
31. Andrade RGd, Chaves OC, Costa DAdS, et al. Overweight in men and women among urban area residents: individual factors and socioeconomic context. Cadernos de Saúde Pública 2015;**31**:148-58
32. Camargos VP, César CC, Caiaffa WT, et al. Imputação múltipla e análise de casos completos em modelos de regressão logística: uma avaliação prática do impacto das perdas em covariáveis. Cad Saúde Pública 2011;**27**:2299-313
33. Ferreira AD, César CC, Malta DC, et al. Validade de estimativas obtidas por inquérito telefônico: comparação entre VIGITEL 2008 e Inquérito Saúde em Beagá. Rev Bras Epidemiol 2011;**14**(1):16-30
34. Gerência de Epidemiologia e Informação Secretaria Municipal de Saúde de Belo Horizonte. Índice de vulnerabilidade à saúde. Secondary Índice de vulnerabilidade à saúde 2003. http://www.pbh.gov.br/smsa/biblioteca/gabinete/risco2003.
35. Friche AA, Xavier CC, Proietti FA, et al. *Saúde urbana em Belo Horizote*. 1 ed, 2015.
36. Friche AA, Diez-Roux AV, Cesar CC, et al. Assessing the Psychometric and Ecometric Properties of Neighborhood Scales in. Journal of urban health : bulletin of the New York Academy of Medicine 2012 doi: 10.1007/s11524-012-9737-z[published Online First: Epub Date]|.
37. Dephmap [program]. 4.06 version. Universiy of London, 2004.
38. Song Y, Knaap G-J. Measuring urban form: Is Portland winning the war on sprawl? Journal of the American Planning Association 2004;**70**(2):210-25
39. Demográfico IC. Resultados do universo. Rio de Janeiro: IBGE 2010
40. Barros AJ, Victora CG. Indicador econômico para o Brasil baseado no censo demográfico de 2000. Revista de Saúde Pública 2005;**39**(4):523-29
41. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health 2006;**60**(4):290-7 doi: 10.1136/jech.2004.029454[published Online First: Epub Date]|.
42. Twisk J. Applied multilevel analysis: a practical guide. 2006: Cambridge University Press, Cambridge, UK.
43. StataCorp L. Stata survey data reference manual, 1985.
44. Sastry N, Pebley AR, Zonta M. Neighborhood definitions and the spatial dimension of daily life in Los Angeles. California Center for Population Research 2002

45. Stewart T, Duncan S, Chaix B, et al. A novel assessment of adolescent mobility: a pilot study. International journal of behavioral nutrition and physical activity 2015;**12**(1):18
46. Hasanzadeh K, Broberg A, Kyttä M. Where is my neighborhood? A dynamic individual-based definition of home ranges and implementation of multiple evaluation criteria. Applied Geography 2017;**84**:1-10
47. Vallée J, Le Roux G, Chaix B, et al. The 'constant size neighbourhood trap'in accessibility and health studies. Urban Studies 2014:0042098014528393

Table 1. Univariate analysis of individual variables by perceived neighbourhood scale: percentages, means and standard deviations.

INDIVIDUAL VARIABLES	Perceived neighbourhood scale (1-4 and %)*				OR (CI 95%) ¹	p
	1 (57.8%)	2 (23.3%)	3 (7.4%)	4 (11.5%)		
Gender (female)	56.4	51.0	45.9	45.9	0.70 (0.58 – 0.83)	<0.001
Employment state (working)	62.0	65.6	69	73.3	1.36 (1.14 - 1.61)	<0.001
Presence of child younger than 10 years (yes)	33.1	31.5	30.6	33.0	0.95 (0.82 - 1.11)	0.540
Number of relatives and friends living in the same neighbourhood (almost all)	2.92	7.07	7.83	13.03	6.30 (4.00 - 9.92)	<0.001
Recognizes most of them people passing by the door of his/her house (yes)	8.0	12.6	13.1	21.3	5.55 (3.04 - 10.11)	<0.001
	Mean (standard deviation)				OR (CI 95%) ¹	p
Age (years)	44.7 (0.35)	44.9 (0.57)	41.1 (0.92)	43.2 (0.74)	0.99 (0.99 - 1.00)	0.060
Socioeconomic position (NEI) **	586.8 (4.06)	601.6 (6.6)	582.8 (11.4)	601.4 (0.0)	1.13 (1.04 - 1.23)	<0.001
Time of residence in the same neighbourhood (years)	14.8 (0.26)	16.8 (0.44)	16.2 (0.71)	16.6 (0.60)	1.01 (1.01 - 1.02)	<0.001

* (1) up to the block or street you live on; (2) within 5 blocks; (3) within ten blocks; and (4) more than ten blocks away; OR – odds ratio; CI 95% – 95% confidence interval; NEI - national economic index; ** Odds ratio calculated based on an increase of 200 NEI points; ¹ estimated by a multilevel ordinal logistic regression model, with reference category being the smaller neighbourhood.

Table 2. Univariate analysis of contextual variables by perceived neighbourhood scale: percentages, means and standard deviations

CONTEXTUAL VARIABLES	NEIGHBORHOOD EXTENSION SCALE (1-4 and %)*				OR (CI 95%) [1]	p
	1 (57.8%)	2 (23.3%)	3 (7.4%)	4 (11.5%)		
Connectivity [2]						
0 to 3	39.91	40.91	41.87	40.37	1.00	
4	23.92	23.42	36.11	25.28	1.04 (0.83 - 1.30)	0.760
5 to 9	37.17	35.63	22.02	34.34	0.85 (0.70 - 1.04)	0.120
	Mean (standard deviation)				OR (CI 95%) [1]	p
Aesthetic quality domain	2.96 (0.03)	3.07 (0.04)	3.06 (0.07)	3.09 (0.04)	1.23 (1.03 - 1.46)	0.020
Walking environment domain	3.20 (0.01)	3.24 (0.02)	3.27 (0.02)	3.28 (0.02)	3.37 (2.09 - 5.44)	<0.001
Violence scale domain	1.90 (0.02)	1.95 (0.03)	1.89 (0.05)	2.00 (0.04)	1.20 (1.03 - 1.40)	0.020
Safety scale domain	2.96 (0.03)	2.93 (0.05)	2.89 (0.09)	2.86 (0.05)	0.92 (0.79 - 1.08)	0.190
Population density (per square kilometre)	12487.35 (791.26)	11740.34 (704.83)	12274.31 (740.63)	12627.22 (865.05)	1.00 (0.99 - 1.01)	0.650

* (1) up to the block or street you live on; (2) within 5 blocks; (3) within ten blocks; and (4) more than ten blocks away; OR – odds ratio; CI 95% – 95% confidence interval. [1] estimated by a multilevel ordinal logistic regression model, with reference category being the smaller neighbourhood; [2] 0 indicates poorly connected streets and 9 indicates heavily connected streets.

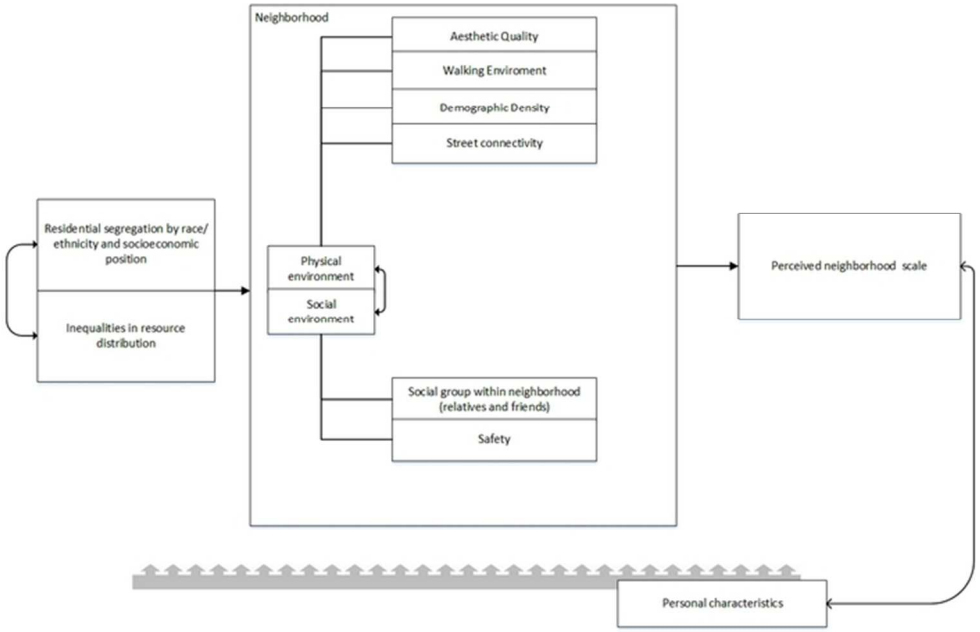
Table 3. Multilevel ordinal logistic regression for the resident perceived neighbourhood scale

Variables	Null model	Contextual variables		Contextual variables + individual variables	
		OR (CI 95%) [1]	p	OR (CI 95%) [1]	p
Aesthetic quality domain		1.21 (0.97 - 1.41)	0.060	1.13 (0.92 - 1.39)	0.230
Walking environment domain		2.96 (1.71 - 5.13)	<0.001	2.22 (1.29 - 3.82)	<0.001
Violence domain		1.35 (1.12 - 1.62)	<0.001	1.23 (1.01 - 1.51)	0.040
Safety domain		0.97 (0.82 - 1.14)	0.710	0.99 (0.83 - 1.19)	0.950
Connectivity [2]					
4		1.00 (0.80 - 1.26)	0.940	1.06 (0.85 - 1.34)	0.570
5 to 9		0.82 (0.67 - 1.01)	0.060	0.89 (0.72 - 1.11)	0.310
INDIVIDUAL					
Age (years)				1.00 (0.99 - 1.00)	0.350
Gender (female)				0.81 (0.68 - 0.96)	0.020
Number of relatives and friends living in the same neighbourhood (almost all)				4.63 (2.84 - 7.57)	<0.001
Recognizes most of the people passing by his/her house (yes)				3.33 (1.72 - 6.25)	<0.001
Employment state (working)				1.26 (1.06 - 1.50)	0.010
Time of residence in the same neighbourhood (years)				1.01 (1.00 - 1.01)	0.130
Presence of child younger than 10 years old (yes)				0.97 (0.81 - 1.16)	0.740
Socioeconomic position**				1.17 (1.06 - 1.29)	<0.001
MODEL INFORMATION					
Variance	0.2567		0.2292		0.2336
MOR	1.62		1.58		1.58
Proportional change in variance	-		10.71		9.00
AIC	8749.26		8668.44		8091.83

OR - odds ratio; CI 95% - 95% confidence interval; NEI - national economic index; MOR - median values of the odds ratio; AIC - Akaike information criterion; ** Odds ratio calculated based on an increase of 200 NEI points; [1] estimated by a multilevel ordinal logistic regression model, with the reference category being the smaller neighbourhood; [2] 0 indicates less connected streets and 9 indicates heavily connected streets.

Figure Legend:

Figure 1- Theoretical model for factors associated with perceived neighbourhood scale



Theoretical Model for perceived neighborhood scale

30x19mm (600 x 600 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation	
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	x
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	x
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	x (pages 3 and 4)
Objectives	3	State specific objectives, including any prespecified hypotheses	x (page 5)
Methods			
Study design	4	Present key elements of study design early in the paper	x (page 5)
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	x (page 5)
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	x (page 5)
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	x (pages 5 to 7)
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Not applicable.
Bias	9	Describe any efforts to address potential sources of bias	Not applicable.
Study size	10	Explain how the study size was arrived at	x (page 5)
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	x (pages 7 and 8)
		(*b*) Describe any methods used to examine subgroups and interactions	Not applicable.
		(*c*) Explain how missing data were addressed	Not applicable.
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	x (page 8)
		(*e*) Describe any sensitivity analyses	Not applicable.
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	x (page 8)
		(b) Give reasons for non-participation at each stage	Not applicable.
		(c) Consider use of a flow diagram	Not applicable.
Descriptive data	14*	(a) Give characteristics of study participants (eg	x (page 8)

		demographic, clinical, social) and information on exposures and potential confounders	
		(b) Indicate number of participants with missing data for each variable of interest	Not applicable.
Outcome data	15*	Report numbers of outcome events or summary measures	x (pages 8 and 9)
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	x
		(*b*) Report category boundaries when continuous variables were categorized	x
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Not applicable.
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	Not applicable.
Discussion			
Key results	18	Summarise key results with reference to study objectives	x (page 9)
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	x (page 11)
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	x (page 11)
Generalisability	21	Discuss the generalisability (external validity) of the study results	x (page 11)
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	x (page 13)

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.