

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Published in final edited form as:

Health Place. 2011 September ; 17(5): 1113–1121. doi:10.1016/j.healthplace.2011.05.011.

The socio-spatial neighborhood estimation method: An approach to operationalizing the neighborhood concept

Malcolm P. Cutchin^a, Karl Eschbach^b, Christine A. Mair^c, Hyunsu Ju^d, and James S. Goodwin^b

^a Department of Allied Health Sciences, University of North Carolina, CB #7122, Bondurant Hall, Suite 2050, Chapel Hill, NC 27599-7122

^b Department of Internal Medicine, Division of Geriatrics, University of Texas Medical Branch, 301 University Blvd, Galveston, TX 77555-0460

^c Department of Sociology and Anthropology, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250-0002

^d Department of Preventive Medicine and Community Health, University of Texas Medical Branch, 301 University Blvd, Galveston, TX 77555-0460

1. Introduction

Research aimed at discovering how the social and physical environment affects health and well-being has increased rapidly during the last decade. A significant proportion of that inquiry focuses on the effects of so-called “neighborhoods.” Recognition of concomitant methodological difficulties in neighborhoods and health research, however, is growing. Further, such methodological difficulties ultimately impact the quality of inferences with regard to health outcomes. One of the most compelling challenges to the quality of inferences from this literature is the definition of the neighborhood as an ecological unit of analysis (Riva et al., 2007). In this paper, we argue that central among those challenges is the development of a methodology to create neighborhood boundaries that define areas in a conceptually valid manner (as opposed to census tracts, for example). Such a method would also help generate more reliable measures of social ecologies, which can then be linked to a host of well-being measures. Moreover, the challenge includes the need to balance the limitations inherent in any such method with the difficulty of use in research. The ultimate goal in addressing these issues in social epidemiology is to improve inferences about the relationship between a very important type of place in human experience—the neighborhood—and various dimensions of health and well-being.

Accordingly, colleagues involved in the Texas City Stress and Health Study developed a new methodology—the Socio-spatial Neighborhood Estimation Method (SNEM). SNEM was designed for creating conceptually-informed neighborhood boundaries. This method became the basis for data collection and, eventually, multi-level modeling of ecological and individual variables. The SNEM approach employs principles of neighborhood boundary definition along with remote (aerial or satellite) imagery to draw boundaries and create neighborhoods for research purposes. Although innovative, the approach is relatively easy to

© 2011 Elsevier Ltd. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

employ by any researcher with access to a basic geographical information system in which to view aerial images of urban space.

In this report, we provide a rationale for this approach and explain its potential significance for the small area effects (on health) literature. We then describe and illustrate the method. To test the value of the approach, we report a two-step “ecometric” (Raudenbush and Sampson, 1999) comparison of ecological measures obtained using the SNEM areas versus the same measures obtained with census block groups and regular grid cells in the study area. Comparisons of intra-class correlation coefficients (ICCs) and establishment of criterion-related validity by way of multi-level models support the conclusion that the areas created with the SNEM approach are more neighborhood-like and provide a better basis for measurement and inferences about neighborhood effects than census boundaries or a regular grid matrix. We conclude with a discussion of the merits and limitations of the SNEM methodology for health research.

2. Background

The rise of social epidemiology and interest in neighborhoods and health was motivated by research on stress as well as the re-emergence of ecological thinking in epidemiology (Berkman and Kawachi, 2000). The synthesis of these two motivating factors indicated that neighborhoods are an important ecological unit that can act as a stressor, or complex of stressors, to affect health. Suggestions of causal links between neighborhoods and health have long extended beyond the stress model to include ideas that neighborhoods might shape behavior that affects health, or that exposures to risk factors in neighborhoods affects health (Berkman and Kawachi, 2000). The ecological dimensions of neighborhoods that could have such affects include the physical (e.g., air, water), built (e.g., housing), services (e.g., education), sociocultural (e.g., ethnic composition), and reputation (e.g., perceptions of the area) (Macintyre and Ellaway, 2003). Macintyre and Ellaway (2003) also emphasized that the differentiation between compositional (population) and contextual (environment) effects is often overstated. Indeed, some causal models are more sensitive to the interplay of neighborhood and personal dimensions as well as the social, environmental, and biological (Daniel et al., 2008). Such model development was driven by findings in the literature that suggested neighborhood characteristics, ranging from sidewalks to complex social relations, are associated with a long list of health risks and outcomes, including various diseases, general measures of mental and physical health, biologic risk factors, health behavior, care utilization, and mortality (Daniel et al., 2008). The social epidemiology literature, therefore, offers a wide range of theory and evidence about the many ways in which neighborhoods are relevant for individual and population health across the life course.

Some researchers tested the effect of defining neighborhoods differently on the relationship between neighborhood attributes and health, concluding that where boundaries are drawn and the size of areas do not substantially affect inferences (Haynes et al., 2008; Jones et al., 2010; Oliver and Hayes, 2007; Reijneveld et al., 2000; Ross et al., 2004; Stafford et al., 2008; Tarkiainen et al., 2010). Others concluded that modes of determining neighborhood boundaries that are sensitive to local context can improve inferences into neighborhood effects on health (Flowerdew et al., 2008; Krieger et al., 2002; Kwan, 2009; Riva et al., 2008; Riva et al., 2009). There are new methods of defining neighborhood areas for health research, such as automated zone design that show potential for improved insights into neighborhoods and health (Cockings and Martin, 2005; Flowerdew et al., 2007; Grady and Enander, 2009). Such operationalizations appear to be less conceptually-informed than is optimal, however.

It is not our purpose to review the burgeoning neighborhoods and health literature, nor is it our intent to review the various issues that challenge this interesting line of investigation into health outcomes. Rather, we focus our attention on one particular issue of empirical, methodological, and theoretical significance: the development of a more refined process of neighborhood operationalization. Central to the enterprise of assessing statistical associations between neighborhoods and health, this operationalization process includes (a) neighborhood conceptualization, (b) determining criteria that reflect the conceptualization, (c) employing those criteria to define areas in residential areas that represent neighborhoods, (d) measuring defined neighborhoods' properties with previously validated instruments, and (e) assessing those neighborhood measures in order to determine if the new operational process holds the potential for inferential gains in neighborhoods and health research. Although until recently the issue of poor operationalization of neighborhoods has been relatively overlooked, this weakness in the literature potentially renders suspect the findings from much of the neighborhoods and health research to date. If neighborhoods continue to be poorly conceptualized and defined, the measures derived to represent them and the inferences drawn from the analysis of those measures also will be problematic and questionable—and by extension, our understanding of important processes linking neighborhood places and health will be weakened. In this section, we lay out the problems and difficulties inherent in this operationalization challenge and provide an argument for our unique methodological approach toward its solution.

2.1 Conceptual issues

Neighborhoods continue to be under-conceptualized in the empirical research on neighborhoods and health (O'Campo, 2003; Diez-Roux, 2007; Riva et al., 2007) and such conceptualization presents a “vexing problem” based in an “elusive consensus” of what a neighborhood is and how it might be best operationalized (Gauvin et al., 2007). There are substantial reasons for researchers to develop a defensible and workable conceptualization. Among those reasons is the need to theorize from empirical findings—an issue that continues to suffer in current research as a result of initial, poor conceptualization and measurement (‘misspecification’) of neighborhoods (Diez-Roux, 2007). Thorough operationalization of neighborhoods should be based in the translation of abstract concepts into concrete and measurable representations (Mueller, 2003). Inferences based on multilevel models using data to represent neighborhood characteristics are of less value if not conceptualized and operationalized rigorously.

There are useful contributions in the literature from which to base a conceptualization of neighborhoods. Among them is the work by Galster (2001), who defined neighborhood as “the bundle of spatially based attributes associated with clusters of residences, sometimes in conjunction with other land uses” (p. 2112). Galster set out ten types of characteristics that make up the neighborhood. Among those are structural (e.g., housing) and infrastructural (e.g., streets) characteristics as well as social, political, economic, environmental and psychological ones. He hastened to note, however, that his conceptualization does not lead to the “Holy Grail” of neighborhood analysis, i.e., a way to bound urban neighborhoods for analysis.

In a similar vein, Chaskin (1997) argued for three essential dimensions of neighborhoods: social, physical (geographical), and experiential. Like Galster, Chaskin questioned the ability to derive a universal definition of neighborhood. Moreover, he suggested that conceptual and operational work on neighborhoods should be based on local contexts and understanding. Research by Grannis (1998, 2009) offered both theoretical and empirical support for his argument that neighborhoods (tertiary communities) are shaped largely by street networks. Larger streets serve as barriers and smaller streets promote cohesion by

facilitating neighborly contact. Grannis' insight was that neighborhoods are formed as much by the larger urban morphology as by internal characteristics.

More directly oriented toward the problem of neighborhoods and health, Sampson, Morenoff and Gannon-Rowley (2002) reviewed the literature on neighborhood effects and conclude that, although the literature was inconsistent in conceptual and empirical approaches as well as findings, four independent and valid classes of neighborhood mechanisms could be discerned. These features and processes set small residential areas apart from one another and include: (a) social ties/interactions, (b) social norms and collective efficacy (i.e., social cohesion), (c) institutional resources, and (d) routine activities (e.g., land use patterns). Sampson et al. (2002) also noted the definition of neighborhood boundaries as an important and needed advance, citing Grannis' work as a fundamental first step.

These applications of social theory to the problem of neighborhoods and health enable at least one synthetic conclusion about neighborhoods that assist with the problem of operationalization. Unlike the inclination to separate the social and the geographical dimensions of neighborhoods (O'Campo, 1993; Gauvin et al., 2007; Spielman and Yoo, 2009), these conceptualizations promote the idea that the social and spatial are joined in the entity we call "neighborhoods." The spatial (e.g., variation in land use and movement between major streets) and social (e.g., interaction and cohesion) processes *influence each other* in an ongoing fashion to create the uniqueness of each neighborhood. The experience of neighborhood influences, then, is reflexively *socio-spatial*, a concept with long-standing currency in the discipline of geography (Soja, 1980).

2.2 Operational issues

If one takes a socio-spatial approach to consider how a neighborhood might be defined and then measured, there are several interrelated operational issues with which to contend. These include the areal size of a neighborhood, the population of a neighborhood, variation within and between neighborhoods, and how to go about determining the boundaries. The primary mode of response to these issues is to ignore them by using available data collected and coded within predetermined (pre-bounded) geographical units. Research on neighborhoods and health is dominated by the use of secondary data generated in this manner, and the most frequent operationalization of the neighborhood in the United States is the census tract.

Although census tract level data are convenient, the census tract itself may comprise an area much larger than what would represent neighborly social ties and interactions, for example. Perhaps more significant is the average population size of 4000 persons per census tract. This population size also represents a group that is relatively large and thus may not reflect the neighborhood concept, particularly in terms of internal social relations and cohesion. Bond Huie and colleagues (2001, 2002) also suggested that units other than the census tract would be more consistent with what constitutes a sociologically meaningful neighborhood: social contact within a relatively small geographic area. Bond Huie et al. (2002) also argue in favor of the methodological benefits of using smaller areas in order to enhance within-area homogeneity and statistical inference. In their research, they used National Health Interview Survey-defined VSAs (very small areas). VSAs are clusters of 4–8 housing units, and while we support the argument of Bond Huie and colleagues for alternatives to census tracts, VSAs are probably too small to be representative of other social and spatial characteristics of neighborhoods. We hasten to note, however, that Bond Huie et al. echoed the implicit argument made by other theorists that neighborhoods conceptually represent places that are distinct from one another and should thus express maximal variation between each other and minimal variation within.

An alternative to census tracts and VSAs could be U.S. census block groups. Yet, those pre-defined areas do not necessarily consider the structural or infrastructural dimensions of a neighborhood. A potential solution to this problem would be the development of a method to construct neighborhoods that are sensitive to local context and based in the conceptual principles of socio-spatial neighborhoods. This potential solution would also need to find a middle ground between the size and population of census tracts and very small areas.

The issues of areal boundaries and size are by no means agreed upon or solved, and they are receiving increased attention. We argue, however, that because the issues noted above remain unresolved, and because methods to address them have rarely been explicit and thorough in the use of neighborhood theory, a need to develop and test alternative approaches to neighborhood construction for health studies exists. The remainder of this article reports the development and assessment of a viable approach to the problem within the context of a social epidemiologic study.

3. The socio-spatial neighborhood estimation method

With these conceptual and operational issues in mind, we endeavored to develop an approach to define (bound) neighborhood areas within a study of stress and health. This section reports the context of the larger study, the goals of the method, the way the method was conducted, and the data and analysis used to evaluate the method.

3.1 Study, place, and goals

The method was devised within the context of the Texas City Stress and Health Study, a social epidemiological study undertaken in a mid-sized city on the gulf coast of Texas. One of the dominant characteristics of Texas City is the large petrochemical complex that borders the residential part of the city (see Figure 1). The study area was carved out of a larger residential area and represented a zone of about 20 square kilometers. The approximate resident population within the study area was 35,000 individuals within approximately 11,000 dwellings. This zone provided a primarily residential area that varied significantly by socio-economic status, ethnic background, and exposure to the hazard. We were interested in examining the relationship of residents to that technological hazard, but we also desired a way to understand the relationship of neighborhood dimensions to stress and health across the varied ethnic population in the city. Developing a plan to implement multilevel modeling techniques was another factor, and thus one important goal was to define enough neighborhoods to model variation along consistent conceptual and operational premises.

3.2 The SNEM approach

The method can be considered a form of qualitative GIS (see Cope and Elwood, 2009). This approach uses imagery to infer qualitative differences between areas—in this case, neighborhoods. Digital orthographic photography quadrants covering the study area were obtained from the Texas Natural Resources Information System and processed in ArcGIS to create one usable image for further analysis. The final image was used to conduct the analysis of the study area and determine boundaries for neighborhoods within that area. ArcGIS was used to conduct the first stages of the analysis. Aerial imagery such as that used in this case allows a viewer to take a bird's eye view of socio-spatial elements of a city and begin to discern patterns that can be used to infer neighborhood boundaries. One of the advantages of the images is the ability to zoom in and out (change the scale of view) in order to see broader patterns as well details that are useful.

Based on the conceptual basis of neighborhoods reviewed above, there were five primary criteria used to discern where boundaries should be drawn to create neighborhoods: (a) street patterns (physical networks that structure social interaction patterns), (b) residential patterns (shifts in housing type and density that signal differences in social class and norms), (c) non-residential land use (barriers to social interaction and signs of resource differences), (d) landforms (potential barriers to interaction), and (e) that neighborhoods should be relatively compact, i.e., not geographically elongated. One analyst (the first author) was responsible for examining the study area image looking for these patterns and features that would indicate where boundaries should be drawn to form neighborhoods.

The method worked in practice through an iteration of photo analysis and on-the-ground examination. Figure 2 illustrates the development of the boundaries for neighborhood 7. The relatively high-resolution imagery allows an analyst to look alternately at larger and smaller (more detailed) portions of the study area. Initial attention was focused on the street grid, and the figures show how larger (primary or secondary) streets are readily identifiable in the image. After checking that interpretation was correct by inspecting other maps, those larger streets became prime candidates for neighborhood boundaries. Other evidence in the images needed to be examined to confirm or deny those preliminary and provisional inferences about boundaries.

Major streets were the primary starting point for determining provisional neighborhood boundaries. Streets were used as primary in the process for two reasons: (a) they were more universal across the study area compared to shifts in residential patterns, commercial land use, and landforms, and (b) other criteria tended to be structured by street patterns; i.e., residential development and land use in particular are dependent on street patterns. Where major streets existed and other criteria did not appear to be present to suggest an alternative boundary to the street basis, the boundaries were relatively easy to draw. For instance, there was no discernable shift in lot size, housing density, housing size, or housing type between neighborhood 7 and 20 (Figure 2; neighborhoods surrounding neighborhood 7 are indicated by their number but their boundaries not shown). Yet there were differences in the residential patterns of lot size and housing type, density, or size in neighborhood 7 and those in numbers 4, 6, and 21. Those differences in residential patterns signaled probable social class differences between the areas that are manifest in the spatial structure of the neighborhoods, and they thereby reinforced the street basis for the boundary. Neighborhood 7 also contained commercial properties along the southern border that can be seen in the image. Such land use most likely served as a buffer between neighborly relations of residents in neighborhoods 7 and 3, and that criterion was used to confirm the boundary originally placed along the major street. Other land uses in the image help to confirm the boundaries suggested initially by the streets. Vacant land in the southwest quadrant and public use in the northwest quadrant indicate boundaries for other neighborhoods. Industrial land use (another buffer) on the eastern side of neighborhood 7 also helped confirm the eastern boundary. Finally, a wide drainage ditch at the far left of the image was used as the western boundary for neighborhood 8.

Confirmation and disconfirmation of boundaries was employed in a second phase of analysis. An additional, essential step in the method was to spend time on the ground, moving through the study area observing the provisional boundary areas and other indicators that could confirm or suggest corrections to the neighborhood boundaries created for the entire study area. After closer inspection by car and by foot, several changes to the boundaries were made. For example, neighborhood 9 was subsumed into neighborhoods 8 and 10. Neighborhood 13 was incorporated into neighborhood 14. Ground-level observation of similarity in lot size, housing type, size, and age helped to make those determinations. Other neighborhood boundaries were moved slightly after observations at ground-level

revealed street widths and traffic patterns suggested that what had seemed to be a major street in the image was not identified properly.

The final set of 48 neighborhoods is shown in Figure 3. Because census tract boundaries are so frequently used to define neighborhoods in health studies, in Figure 3 we also show those boundaries to indicate comparative size and shape. Much smaller than census tracts, the neighborhoods contained, on average, approximately 230 residential units and 730 residents. Also apparent on that map are non-residential areas of significant size that were left out of bounded neighborhoods, primarily because they probably act as buffers to interaction among residents on different sides of those spaces.

3.3 Assessment of the method

The first objective of assessment was to determine if the SNEM method resulted in socio-spatial units that have similar characteristics across the particular neighborhood but that are relatively distinct from other neighborhoods in the urban area. Following the arguments of “ecometrics” (Raudenbush and Sampson, 1999), intra-class correlation coefficients (ICCs) of ecological measures that provide a ratio of intra-area to intra- and inter-area variation offer an appropriate way to assess and then compare SNEM-generated neighborhood properties to neighborhoods generated with other possible approaches. The higher the ICC on a given indicator, the better the method’s ability to generate areas that represent the intra-area homogeneity and inter-area variation aspects of the neighborhood concept.

The second, and more instrumental objective for determining if SNEM could enhance inferences about neighborhood effects on health, was to compare the performance of SNEM-generated measures in multi-level models with measures collected through other neighborhood types. This step was included in order to provide essential criterion-related validity evidence for the SNEM approach. The comparison was based on whether the association of SNEM neighborhood properties with an individual-level health outcome measure (criterion) meets or exceeds that of associations of the criterion with indicator measures estimated with other neighborhood types. The second analysis aimed to determine the method’s validity at producing neighborhood-level health predictors vis-à-vis other neighborhood definition approaches, and it thus aimed to assess whether inferential ability regarding health is thereby improved by the SNEM approach.

3.3.1 Data collection—Data came from the Texas City Stress and Health Study, part of a larger Center for Population Health and Health Disparities focusing on health of Hispanics. The original purpose of the study was to investigate various patterns of ecological, psychosocial, and biological risk factors and their association with stress and health outcomes in Texas City (Cutchin, 2007; Cutchin et al., 2008; Mair et al., 2011; Peek et al., 2008; Peek et al., 2009a; Peek et al., 2009b; Stowe et al., 2010a; Stowe et al., 2010b). There was a particular emphasis on Mexican-Americans because little is known about that group vis-à-vis others within those larger patterns. Working neighborhood by neighborhood, the study began with an exhaustive listing of housing units in the study area. From this listing, housing units were classified as Hispanic or other households. Hispanic households had at least one adult who self-identified as Hispanic. The households were divided into three strata: (a) one in eight non-Hispanic white or non-Hispanic black households were selected, and one adult 25+ was randomly selected for interview, (b) one adult aged 25–64 was randomly selected from all Hispanic households with no older adults (65+); and (c) all Hispanic adults aged 65 and over were selected for interview. Selected and consenting residents were interviewed in their homes. The interview response rate was 80% producing a final sample of 2706 respondents. The baseline survey instrument contains scales and items measuring a wide array of demographic, behavioral, social, and health indicators. The

institutional review board at the University of Texas Medical Branch approved the study protocol, and informed consent was obtained from all participants.

3.3.2 Measures—A selection of subjective neighborhood and objective neighborhood measures were used (see Table 1). The subjective measures were collected with the perceived neighborhood scale (Martinez et al., 2002). That instrument produces individual-level perceptions of neighborhood quality and social relations. Besides a comprehensive score, 4 subscale scores represent the following concepts: (a) perceived crime, (b) neighborhood satisfaction, (c) social embeddedness (ties/interaction), and (d) sense of community (social cohesion). Martinez et al. (2002) reported reliability coefficients above .80 for all scales. When implementing the instrument in the study area, respondents were asked to rate the area “within a few blocks of their house.” For the data collected in Texas City, Cronbach’s alphas for all five scores ranged between .80 and .91.

Objective measures of neighborhood quality also were collected throughout the study area. Observations of the physical and social environment were collected for all 586 census blocks in the study area using a modified version of an instrument developed and validated by Caughy et al. (2001). That instrument was designed for use by a trained observer who walks city streets and rates various aspects of the urban environment. One block observer was trained in several non-study areas and his work checked and verified. The observer then rated the study area blocks, and again, his work was checked and verified for accuracy by two investigators. Factor analysis, based on 537 blocks with complete data provided two variables for use in this analysis. One was a dimension we named “Territoriality” indicated by the presence of neighborhood signs, amount of graffiti, and proportion of residences with yards. A second factor, “Residential Environmental Quality,” was indicated by four measures: overall condition of most residential units, the type (quality) of building material used in residences, the overall condition of resident-kept grounds, and the amount of litter seen in the block. Factor scores were used for both of these variables in the following analysis.

3.3.3 Analysis—We employed two other modes of neighborhood definition with which to compare SNEM neighborhoods. One was census block groups as neighborhoods, of which there were 31 in the study area (Figure 4).¹ The other was a regular grid of 48 areas placed arbitrarily over the study to create spatial units within which individual residences were aggregated (Figure 5). Grid-based data have been used for years in the Nordic countries for the organization, release, and analysis of statistical data (Tammilehto-Luode et al., 2000). The grid was created using a formula that determined what shape and size of grid cells would allow coverage of the study area with 48 units to match the 48 SNEM units. Although the cells are not perfectly square, they were compact enough to fairly match the average SNEM unit.

The econometric approach to analysis used a linear mixed model to calculate ICCs for all measures across the three neighborhood types. Mixed models partitioned the variance in objective and subjective neighborhood measures into individual- and neighborhood-level components. Unconditional means models containing only a neighborhood-level random intercept were used to estimate the ICCs, which were calculated as the ratio of the neighborhood-level variance component to the total variance. The ICC represents the ratio

¹U. S. Census block groups compose census tracts and typically contain a population between 600 and 3,000 persons (Iceland and Steinmetz, 2003). In the study area in Texas City, the average was approximately 1,150 people per block group. Boundaries of both census tracts and block groups were originally drawn up by local committees in metropolitan areas concerned with collecting administrative and health data for planning purposes (Krieger, 2006). The specific practices used to create the original boundaries are unclear.

of between-neighborhood variance to the sum of between- and within-neighborhood variance for each of the subjective and objective measures. Our hypothesis was that SNEM neighborhood measures would have overall higher ICCs, indicating the method results in socio-spatial areas that are more reflective of the conceptualization of neighborhood.

For the determination of criterion-related validity, we estimated multi-level models using the same subjective and objective neighborhood measures. The models were estimated in STATA (v. 11.1) using XTMIXED with random effects for the neighborhood units. Each model included the neighborhood-level independent variable (e.g., neighborhood aggregate mean of individual scores on perceived crime) and individual-level independent variables of age (continuous), gender, marital status (married or not), education (less than high school, high school graduate, some college, college graduate), race/ethnicity (non-Hispanic black, non-Hispanic white, native Hispanic, immigrant Hispanic), and the individual score on the subscale (or the factor score for the block where the individual lived). The SF-36 v.2 (Ware, 2000) summary scores for the physical and mental components of health-related quality of life were used as dependent variables in the models. The SF-36 is widely used in population health and clinical studies, and it contains eight dimensions: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health. Reliability and validity across population groups has been well established. The resulting z-score coefficients for the neighborhood-level variables were then compared across different neighborhood types to assess whether the SNEM approach led to improved inferences about neighborhoods and health relationships. We hypothesized that SNEM neighborhood measures would exhibit a pattern of more statistically significant relationships with the health-related quality of life scores.

4. Findings

Table 2 shows the ICCs estimated for all measures and neighborhood types. The SNEM neighborhoods produced better ICCs for subjective measures in all cases except for the social embeddedness score, where block groups performed better. Objective measures results also indicate good performance by SNEM neighborhood units relative to the other two neighborhood types.

Table 3 reports the results of the criterion-related validity assessment. The boldface z-scores indicate statistically significant relationships between the neighborhood-level variables and the health-related quality of life scores from the SF-36. Noteworthy are the four significant findings for the SF-36 physical component score using the SNEM-based independent variables (perceived crime, neighborhood satisfaction, sense of community, and residential environmental quality). Block group and grid cell-level independent variables were not significantly associated with the physical component score. When modeling the SF-36 mental component score, block group and grid cell-level measures of sense of community were significant, but while the same measure using SNEM neighborhoods was not significant, it was close (1.85) to the threshold of 1.96. Because of the number of significant effects discovered that would have been undetected using measures created with the other neighborhood types, these results indicate the potential utility of the SNEM approach.

5. Discussion and conclusion

The ICC comparison supports our hypothesis that the SNEM approach results in socio-spatial areas that are more reflective of the neighborhood concept. We should note, however, that the geographical basis for the neighborhood perception items—asking respondents to rate an area a few blocks around their home—probably created some error in those aggregated measures. However, the fact that our SNEM ICCs on the two objective measures

were much larger than those for the other units of analysis is noteworthy. Because the qualitative assessment of residential differences and associated social class differences is an important feature of the method, we should be able to discern more abrupt changes in objective neighborhood characteristics. The relatively large ICCs from objective measures support the conclusion that the method works particularly well to demarcate the space of neighborhoods that are more distinct in a socio-structural way compared to the space of those we obtain via pre-determined boundaries.

Our additional assessment and findings of the relative criterion-related validity of the SNEM approach were perhaps more important. That evidence suggests that the method is more valid than common alternative approaches for drawing inferences about neighborhood and health relationships. Based on our findings, SNEM measures of neighborhood characteristics appear to be more frequently related to health-related quality of life measures, especially for the physical health dimension of the SF-36. If we had used block groups or grid cells as proxies for neighborhoods, we would have in effect committed Type II errors in drawing inferences about the significant relationships between our subjective and objective measures and health-related quality of life. By assessing SNEM-based measures against the same measures collected with different approaches to neighborhood operationalization, we have additional evidence that the new neighborhood operational process holds potential for inferential gains in neighborhoods and health research.

The neighborhoods we generated using the SNEM approach are smaller than those usually employed in neighborhoods effects studies. As Haynes et al. (2007) have pointed out, studies using smaller areas tend to find higher ICCs because of lower intra-area variation. Therefore, our ICCs, especially those for the objective measures, may be advantaged by the small size of our created neighborhoods. We note, however, that the block groups in our study are not that much larger than our neighborhoods and the grid cells are very close to the same size. Yet, the question of the modifiable areal unit problem (MAUP) remains—the documented effect that size of the spatial unit containing aggregated data will influence the correlations within it. The problem has two dimensions, scale and zoning (Fotheringham et al., 2000). That is, both the size of the area and the shape (or location) of its boundaries may bias statistical estimates. Further, as of yet there is no good solution to the problem (Fotheringham et al., 2000).

We reiterate that there are other approaches to generating neighborhoods for health research. One type is automated zone design, a general quantitative approach of finding optimal boundaries for small areas based on the distribution of underlying data (Haynes et al., 2007). Although varied in specifics of how this general type of approach is employed, researchers pre-specify criteria for how the zone design models will operate and thus express some conceptualization of what neighborhood attributes may be relevant to health. Nonetheless, it appears that in most uses the selection of those criteria are not as strongly influenced by theory as the SNEM approach. In the SNEM approach, the method of neighborhood definition flows from a socio-spatial theorization of neighborhood processes—with the understanding that, as with automated zone design, there is an inability to fully represent the complexity of neighborhood processes in operationalization. Moreover, it might be possible to automate at least part of the SNEM approach, although we argue that observation on the ground would remain necessary to confirm or adjust boundaries.

Another approach is to focus on social networks, agency, and agent-based modeling (Entwisle, 2007). Those such approaches allow behavior to enter the frame more clearly than our approach or that of automated zone design, but agent-based modeling also tends to overlook some of the fundamental, structural features of neighborhoods suggested by SNEM, such as residential social class opportunities in the housing market and land use

patterns (the attention to variation in lot size, housing type, housing size, and commercial and traffic buffers). We suggest that although the ‘fit’ between structural patterns and social relations in urban space is not a perfect one, the compromise made with the SNEM approach is to use the probability that patterns indicate relations and provide a method that is comparatively easy for researchers use. Moreover, as a socio-spatial approach, SNEM has a strong conceptual basis that interweaves social and spatial processes and is sensitive to details of local context.

The socio-spatial neighborhood estimation method and our assessment of it have limitations. Only one analyst implemented the method (and he also developed the method). A team approach to determine neighborhood boundaries using imagery would most likely enhance the validity of neighborhood boundaries. Because the method is largely an inductive one similar to qualitative analysis of text, it is unclear if an inter-rater reliability process would be more valuable than discussion and consensus building to decide boundaries. Nonetheless, there is a need for replication of the SNEM approach and the results we report here. Use of other instruments that capture neighborhood compositional and contextual information would enhance assessment. Another limitation is that only one mid-sized city was used to develop and demonstrate the method. While we believe the method has transferability to other types of urban environments, this has yet to be demonstrated. We see no reason to suggest the method would not work in other cities, but we also note that the SNEM approach might find utmost utility in studies of relatively smaller geographical areas and where primary data collection is used.

Those limitations noted, we conclude with the suggestion that the socio-spatial neighborhood estimation method shows promise for use in studies of neighborhood effects. The method incorporates conceptual bases of neighborhoods that are both social and spatial in order to better operationalize their measurement and enhance inferences about neighborhood effects. Research design in this area continues to struggle with both conceptual and measurement issues, and the SNEM offers a useable and logical way to move forward.

Acknowledgments

The authors wish to acknowledge the contributions of Peter Dana and Joseph Forrest on GIS work related to this project, including imagery, maps, geocoding, and analysis. Tasanee Walsh also contributed earlier analysis that led to the work reported here. Steve Owen added statistical suggestions about neighborhoods and sampling as well as about validation. This work was supported by Grant P50 CA10563 from the National Cancer Institute which funded the UTMB Center for Population Health and Health Disparities as well as the Texas City Stress and Health Study.

References

- Berkman, LF.; Kawachi, I. A historical framework for social epidemiology. In: Berkman, L.; Kawachi, I., editors. *Social Epidemiology*. Oxford University Press; Oxford: 2000. p. 3-12.
- Bond Huie SA. The concept of neighborhood in health and mortality research. *Sociological Spectrum*. 2001; 21:341–358.
- Bond Huie SA, Hummer RA, Rogers RG. Individual and contextual risks of death among race and ethnic groups in the United States. *Journal of Health and Social Behavior*. 2002; 43:359–381. [PubMed: 12467258]
- Caughy MO, O’Campo PJ, Patterson J. A brief observational measure for urban neighborhoods. *Health & Place*. 2001; 7:225–236. [PubMed: 11439257]
- Chaskin RJ. Perspectives on neighborhood and community: a review of the literature. *Social Service Review*. 1997; 71:521–547.
- Cockings S, Martin D. Zone design for environment and health studies using pre-aggregated data. *Social Science & Medicine*. 2005; 60:2729–2742. [PubMed: 15820583]

- Cope, M.; Elwood, S. *Qualitative GIS: A Mixed Methods Approach*. Sage; Thousand Oaks, CA: 2009.
- Cutchin MP. The need for the “new health geography” in epidemiologic studies of environment and health. *Health and Place*. 2007; 13:725–742. [PubMed: 17208033]
- Cutchin MP, Martin KR, Owen SV, Goodwin JS. Concerns about petrochemical health risk before and after a refinery explosion. *Risk Analysis*. 2008; 28:589–601. [PubMed: 18643817]
- Daniel M, Moore S, Kestens Y. Framing the biosocial pathways underling assumptions between place and cardiometabolic disease. *Health & Place*. 2008; 14:117–132. [PubMed: 17590377]
- Diez-Roux A. Neighborhoods and health: where are we and where do we go from here? *Revue d'Epidémiologie et de Santé Publique*. 2007; 55:13–21.
- Entwisle B. Putting people into place. *Demography*. 2007; 44:687–703. [PubMed: 18232206]
- Flowerdew R, Feng Z, Manley DJ. Constructing data zones for Scottish Neighbourhood Statistics. *Computers, Environment and Urban Systems*. 2007; 31:76–90.
- Flowerdew R, Manley DJ, Sabel CE. Neighbourhood effects on health: does it matter where you draw the boundaries? *Social Science & Medicine*. 2008; 66:1241–1255. [PubMed: 18177988]
- Fotheringham, AS.; Brunsdon, C.; Charlton, M. *Quantitative Geography: Perspectives on Spatial Data Analysis*. Sage; Thousand Oaks, CA: 2000.
- Galster G. On the nature of neighbourhood. *Urban Studies*. 2001; 38:2111–2124.
- Gauvin L, Robitaille E, Riva M, McLaren L, Dassa C, Potvin L. Conceptualizing and operationalizing neighborhoods: the conundrum of identifying territorial units. *Canadian Journal of Public Health*. 2007; 98:S18–S26.
- Grady S, Enander H. Geographic analysis of low birthweight and infant mortality in Michigan using automated zone design methodology. *International Journal of Health Geographics*. 2009; 8:10. [PubMed: 19224644]
- Grannis R. The importance of trivial streets: residential streets and residential segregation. *American Journal of Sociology*. 1998; 103:1530–1564.
- Grannis, R. *From the Ground Up: Translating Geography into Community Through Neighbor Networks*. Princeton University Press; Princeton, NJ: 2009.
- Haynes R, Jones A, Reading R, Daras K, Emond A. Neighbourhood variations in child accidents and related child and maternal characteristics: does area definition make a difference? *Health and Place*. 2008; 14:693–701. [PubMed: 18166497]
- Haynes R, Konstantinos D, Reading R, Jones A. Modifiable neighborhood units, zone design and residents' perceptions. *Health & Place*. 2007; 13:812–825. [PubMed: 17369075]
- Iceland, J.; Steinmetz, E. The effects of using census block groups instead of census tracts when examining residential housing patterns. US Census Bureau working paper. 2003 July. http://www.census.gov/hhes/www/housing/resseg/pdf/unit_of_analysis.pdf
- Jones AP, van Sluijs EMF, Ness AR, Haynes R, Riddoch CJ. Physical activity in children: Does how we define neighbourhood matter? *Health & Place*. 2010; 16:236–241. [PubMed: 19906555]
- Krieger N. A century of census tracts: health & the body politic (1906–2006). *Journal of Urban Health: Bulletin of the New York Academy of Medicine*. 2006; 83:355–361. [PubMed: 16739037]
- Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: Does the choice of area-based measure and geographic level matter? The Public Health Disparities Geocoding Project. *American Journal of Epidemiology*. 2002; 156:471–482. [PubMed: 12196317]
- Kwan MP. From place-based to people-based exposure measures. *Social Science & Medicine*. 2008; 69:1311–1313. [PubMed: 19665828]
- Macintyre, S.; Ellaway, A. Neighborhoods and health: an overview. In: Kawachi, I.; Berkman, L., editors. *Neighborhoods and Health*. Oxford University Press; Oxford: 2003. p. 20–42.
- Mair CA, Cutchin MP, Peek MK. Allostatic load in an environmental riskscape: The role of stressors and gender. *Health & Place*. 2011; 10.1016/j.healthplace.2011.03.009
- Martinez ML, Black M, Starr RH. Factorial structure of the Perceived Neighborhood Scale (PNS): a test of longitudinal invariance. *Journal of Community Psychology*. 2002; 30:23–43.
- Mueller, CW. Conceptualization, operationalization, and measurement. In: Lewis-Black, MS.; Bryman, A.; Liao, TF., editors. *The Sage Encyclopedia of Social Science Research Methods*.

Sage; Thousand Oaks, CA: 2003 Apr 27. 2010,
http://www.sage-ereference.com/socialscience/Article_n150.html

- O'Campo P. Invited commentary: advancing theory and methods for multilevel models of residential neighborhoods and health. *American Journal of Epidemiology*. 2003; 157:9–13. [PubMed: 12505885]
- Oliver L, Hayes MV. Does choice of spatial unit matter for estimating small area disparities in health and place effects in the Vancouver census metropolitan area? *Canadian Journal of Public Health*. 2007; 98:S27–S34.
- Peek MK, Cutchin MP, Freeman DH, Perez NA, Goodwin JS. Perceived health change in the aftermath of a petrochemical accident: an examination of pre-accident, within-accident, and post-accident variables. *Journal of Epidemiology and Community Health*. 2008; 62:106–112.
- Peek MK, Cutchin MP, Freeman D, Stowe RP, Goodwin JS. Environmental hazards and stress: evidence from the Texas City Stress and Health Study. *Journal of Epidemiology and Community Health*. 2009a; 63:792–798. [PubMed: 19282316]
- Peek MK, Cutchin MP, Salinas JJ, Sheffield KM, Eschbach K, Stowe RP, Goodwin JS. Allostatic load among non-Hispanic whites, non-Hispanic blacks, and people of Mexican origin: effects of ethnicity, nativity, and acculturation. *American Journal of Public Health*. 2009b; 99:1–7.
- Raudenbush SW, Sampson RJ. Ecometrics: toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology*. 1999; 29:1–41.
- Reijneveld SA, Verheij RA, de Bakker DH. The impact of area deprivation on differences in health: does the choice of the geographical classification matter? *Journal of Epidemiology and Community Health*. 2000; 54:306–313. [PubMed: 10827914]
- Riva M, Gauvin L, Barnett TA. Toward the next generation of research into small area effects on health: a synthesis of multilevel investigations published since July 1998. *Journal of Epidemiology and Community Health*. 2007; 61:853–861. [PubMed: 17873220]
- Riva M, Apparicio P, Gauvin L, Brodeur JM. Establishing the soundness of administrative spatial units for operationalising the active living potential of residential environments: an exemplar for designing optimal zones. *International Journal of Health Geographics*. 2008; 7:1–13. [PubMed: 18190678]
- Riva M, Gauvin L, Apparicio P, Brodeur J. Disentangling the relative influence of built and socioeconomic environments on walking: the contribution of areas homogeneous along exposures of interest. *Social Science & Medicine*. 2009; 69:1296–1305. [PubMed: 19733426]
- Ross NA, Tremblay S, Graham K. Neighbourhood influences on health in Montreal, Canada. *Social Science & Medicine*. 2004; 59:1485–1494. [PubMed: 15246176]
- Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing “neighborhood effects”: social processes and new directions in research. *Annual Review of Sociology*. 2002; 28:443–478.
- Soja EW. The socio-spatial dialectic. *Annals of the Association of American Geographers*. 1980; 70:207–225.
- Spielman SE, Yoo E. The spatial dimensions of neighborhood effects. *Social Science & Medicine*. 2009; 68:1098–1105. [PubMed: 19167802]
- Stafford M, Duke-Williams O, Shelton N. Small area inequalities in health: are we underestimating them? *Social Science & Medicine*. 2008; 67:891–899. [PubMed: 18599174]
- Stowe RP, Peek MK, Cutchin MP, Goodwin JS. Plasma cytokine levels in a population-based study: relation to age and ethnicity. *Journals of Gerontology Series A—Biological Sciences and Medical Sciences*. 2010a; 65:429–433.
- Stowe RP, Peek MK, Perez NA, Yetman DL, Cutchin MP, Goodwin JS. Herpesvirus reactivation and socioeconomic position: a community-based study. *Journal of Epidemiology and Community Health*. 2010b; 64:666–671. [PubMed: 19825788]
- Tammilehto-Luode, M.; Rogstadt, L.; Backer, L. GRID data and area delimitation by definition. Statistical Commission and Economic Commission for Europe, Working Paper No. 24. 2000 March 8. 2011, <http://www.unece.org/stats/documents/2000/04/gis/24.e.pdf>

- Tarkiainen L, Martikainen P, Laaksonen M, Leyland AH. Comparing the effects of neighbourhood characteristics on all-cause mortality using two hierarchical areal units in the capital region of Helsinki. *Health & Place*. 2010; 16:409–412. [PubMed: 19897398]
- Ware JE. SF-36 health survey update. *Spine*. 2000; 25:3130–3139. [PubMed: 11124729]



Figure 1.
Texas City Stress and Health Study Area



Figure 2.
Showing neighborhood no. 7's boundaries based on analytical criteria.



Figure 3.
Study area SNEM-generated neighborhoods and census tracts.

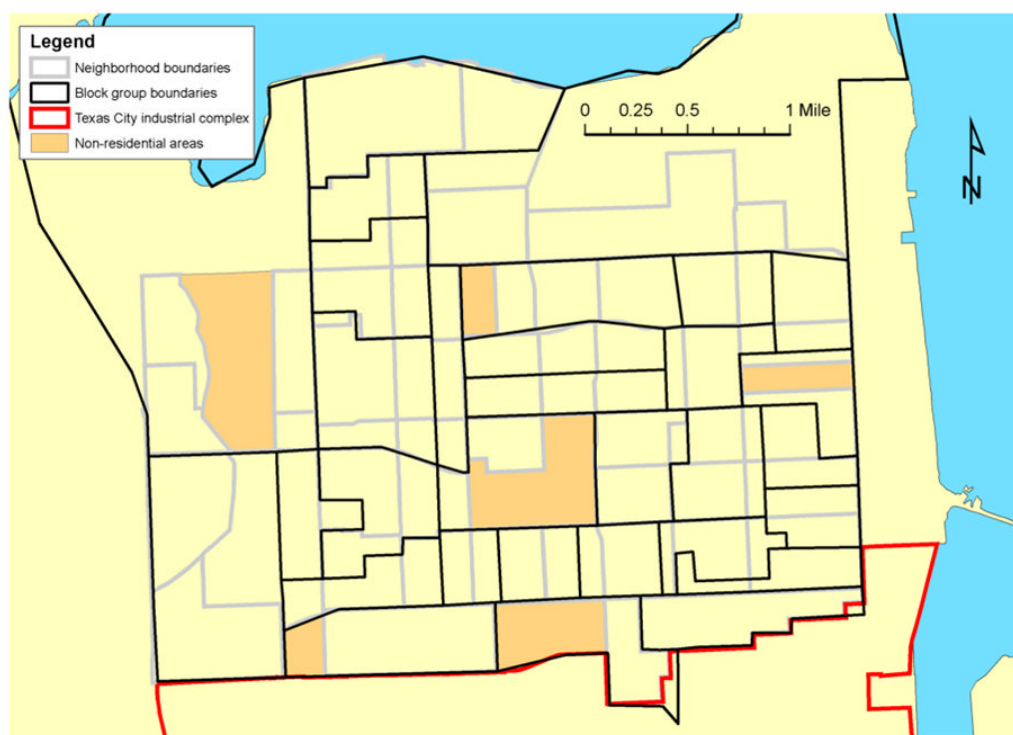


Figure 4.
Study area census block groups and SNEM-generated neighborhoods.

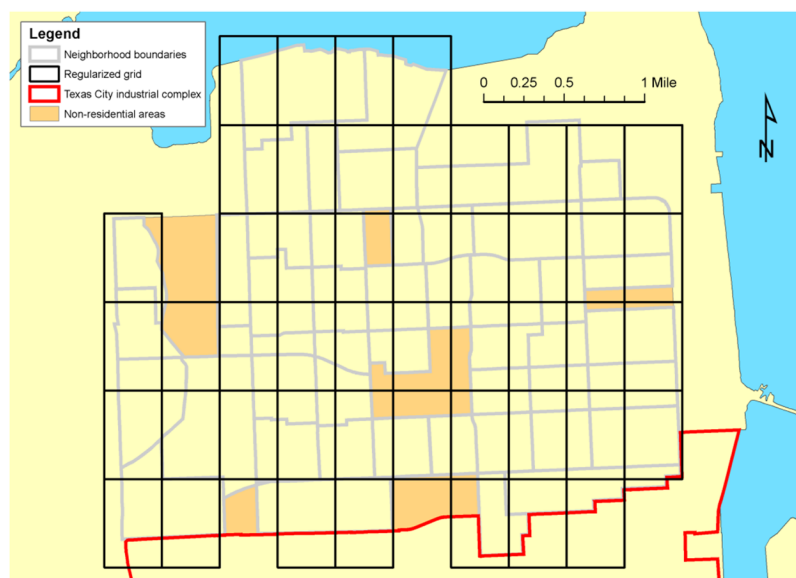


Figure 5.
Uniform grid neighborhoods and SNEM-generated neighborhoods.

Table 1**Subjective and Objective Neighborhood Measures**

Subjective Neighborhood Measures (Perceived Neighborhood Scale, Martinez et al., 2002)	
(a) Perceived Crime	<i>Response Scale (1–5, “strongly agree” to “strongly disagree”)</i>
	1 There are troublemakers hanging around in my neighborhood.
	2 There is public drinking in my neighborhood.
	3 There is open drug abuse or drug dealing in my neighborhood.
	4 It is safe to walk along in my neighborhood at night.
	5 Some of my friends and relatives won’t visit me because they don’t feel safe in my neighborhood.
	6 People are afraid of being robbed in my neighborhood.
(b) Neighborhood Satisfaction	<i>Response Scale (1–5, “strongly agree” to “strongly disagree”)</i>
	7 My neighborhood is a good place to live.
	8 My neighborhood has been getting worse recently.
	9 The buildings and yards in my neighborhood are really run down.
	10 If I could I would move out of my neighborhood.
	11 There is a good place for children to play in my neighborhood.
	12 My neighborhood is a good place to raise a family.
(c) Social Embeddedness	<i>Response Scale (1–5, “very likely” to “very unlikely”)</i>
	13 How likely is it that you would help a neighbor by watching their place if they’re away or taking care of their child if they are sick?
	How likely is it that a neighbor...
	14 Would loan you a few dollars if you needed it in an emergency?
	15 Would ask you to borrow some good if he/she needed some in an emergency?
	16 Would watch your place if you’re away or take care of your child when you’re sick?
	Response Scale (1–5, “very often (daily)” to “never”)
	How often do you...
	17 Greet your neighbors when you see them?
	18 Have casual visits with neighbors either in their home or in yours?
	19 Go to activities like church fairs, neighborhood meetings, or sports events?
	20 Exchange/share child care?
	21 Talk to your neighbors?
(d) Sense of Community	<i>Response Scale (1–5, “strongly agree” to “strongly disagree”)</i>
	22 Among my neighbors there are people I can rely on.
	23 People in my neighborhood trust each other.
	24 I feel I belong in my neighborhood.
	25 I care about what my neighbors think of how I dress and how my house looks.
	26 I have a close relationship with some of my neighbors.
	27 People in my neighborhood are usually warm and friendly.
	28 In my neighborhood people help each other out.
Objective Neighborhood Measures (adapted from Caughy et al., 2001)	

Subjective Neighborhood Measures (Perceived Neighborhood Scale, Martinez et al., 2002)

(a) Territoriality (factor score)	1	Signs denoting neighborhood name (<i>0–1, no or yes</i>)
	2	Proportion of residences with front yards (<i>1–4, none to more than half</i>)
	3	Amount of graffiti (<i>1–4, considerable amount to none</i>)
(b) Residential Environmental Quality (factor score)	1	Overall condition of most residential units (<i>1–4, poor to excellent</i>)
	2	Type (quality) of building material used in residences (<i>1–6 with proportion of wood to proportion brick or stone [and other materials in between]</i>)
	3	Overall condition of resident-kept grounds (<i>1–4, poor to excellent</i>)
	4	Amount of litter seen in the block (<i>1–4, considerable amount to none</i>)

Table 2

Intra-class correlation coefficients of subjective and objective measures using SNEM neighborhood (NH), block group (BG) and grid units of analysis

Subjective Measures	NH ICC	BG ICC	Grid ICC
Perceived crime	0.276	0.227	0.221
Satisfaction	0.228	0.185	0.168
Social embeddedness	0.043	0.060	0.034
Sense of community	0.060	0.044	0.055
Perceived NH Scale	0.192	0.164	0.158
Objective Measures			
Territoriality	0.906	0.347	0.573
Residential environmental quality	0.711	0.556	0.600

Table 3

Z-score coefficients from multilevel models.*

<i>SF-36 Physical Component Score</i>			
	<u>SNEM (n = 48)</u>	<u>Block Group (n = 31)</u>	<u>Grid (n = 48)</u>
Perceived crime (n = 2249)	2.22	0.90	1.16
Satisfaction (n = 2351)	3.44	1.66	1.87
Social embeddedness (n = 2236)	1.32	1.53	0.30
Sense of community (n = 2220)	3.22	1.44	1.76
Territoriality (n = 2334)	-0.37	-0.44	-0.13
Residential environmental quality (n = 2317)	2.67	0.63	1.61
<i>SF-36 Mental Component Score</i>			
	<u>SNEM</u>	<u>Block Group</u>	<u>Grid</u>
Perceived crime (n = 2249)	-0.57	-0.53	-0.01
Satisfaction (n = 2351)	0.88	-0.25	-0.07
Social embeddedness (n = 2236)	-0.23	1.40	0.51
Sense of community (n = 2220)	1.85	2.41	2.11
Territoriality (n = 2334)	-0.76	1.21	-1.45
Residential environmental quality (n = 2317)	0.51	0.83	1.19

* Estimated in STATA using XTMIXED with random effects for areal units. Each z-score was estimated with one model using the neighborhood-level independent variable and including individual-level independent variables of age (continuous), gender, marital status (married or not), education (less than high school, high school graduate, some college, college graduate), race/ethnicity (non-Hispanic black, non-Hispanic white, native Hispanic, immigrant Hispanic), and the individual score on the subscale (or the factor score for the block where the individual lived). Bold z-scores are significant at the $p \leq .05$ level. Sample sizes for models vary based on areal units (noted in column headers) and individuals without missing data (noted in row labels).