Defining Neighborhood Boundaries for Urban Health Research

Linda Weiss, PhD, Danielle Ompad, PhD, Sandro Galea, MD, DrPH, David Vlahov, PhD

Abstract: The body of literature exploring neighborhood effects on health has increased rapidly in recent years, yet a number of methodologic concerns remain, including preferred methods for identification and delineation of study neighborhoods. In research combining census or other publicly available data with surveys of residents and/or street-level observations, questions regarding neighborhood definition take on added significance. Neighborhoods must be identified and delineated in such a way as to optimize quality and availability of data from each of these sources.

IMPACT (Inner-City Mental Health Study Predicting HIV/AIDS, Club and Other Drug Transitions), a multilevel study examining associations among features of the urban environment and mental health, drug use, and sexual behavior, utilized a multistep neighborhood definition process including development of census block group maps, review of land use and census tract data, and field visits and observation in each of the targeted communities. Field observations were guided by a preidentified list of environmental features focused on the potential for recruitment (e.g., pedestrian volume), characteristics commonly used to define neighborhood boundaries (e.g., obstructions to pedestrian traffic, changes in land use), and characteristics that have been associated in the literature with health behaviors and health outcomes (such as housing type and maintenance and use of open spaces). This process, implemented in February through July 2005, proved feasible and offered the opportunity to identify neighborhoods appropriate to study objectives and to collect descriptive information that can be used as a context for understanding study results.

Introduction

The body of literature exploring neighborhood effects on health has increased rapidly in recent years. Social, institutional, and physical characteristics of neighborhoods are being linked to a variety of health behaviors and outcomes including substance abuse, asthma, cardiovascular disease, birth outcomes, respiratory infections, sexually transmitted diseases, cancer, and all-cause mortality. However, including specification of social, institutional, and physical characteristics of neighborhoods that should be included in such research, appropriate methods for collecting neighborhood data (e.g., resident surveys, secondary source data, and/or direct, systematic observation), and guidelines for identification and delineation of study neighborhoods.

Interest in neighborhoods and the impact they have on residents has spanned decades and disciplines, including sociology, child development, and public health. There is a general, if not unanimous, consensus within the literature that neighborhood refers to a geographic unit of limited size, with relative homogeneity in housing and population, as well as some level of social interaction and symbolic significance to residents. The subjectivity of neighborhood boundaries, demonstrated repeatedly in empirical research, is also widely accepted. Social connections, common use of public facilities (e.g., schools, shopping areas), and physical barriers (e.g., major thoroughfares) may contribute to overlap in residents’ neighborhood definitions, but their perceptions are also affected by individual characteristics, such as gender, age, and daily activities. Particularly in cities, where local travel is easy and frequent, neighborhood boundaries are likely to be malleable.

Given the difficulties inherent in neighborhood delineation, public health researchers have most often opted for predefined boundaries consistent with socio-
demographic and health data available from secondary sources (e.g., census tracts in the United States, electoral wards in the United Kingdom). Predefined boundaries are easily identified, replicable, and obviously allow for the use of secondary source data. Their disadvantage rests in possible discrepancies with contemporary settlement patterns and resident perceptions of neighborhood boundaries. Local-level variability may be obscured within these preselected units, particularly if they are too large. In addition, exclusive reliance on secondary source boundaries may be inefficient for studies involving primary data collection, as it may be impossible to discern, prior to the start of data collection, if factors of interest are present in the selected geographic areas. Despite these limitations, there is little in the public health literature suggesting that alternative methods for delineating neighborhood boundaries have been attempted. Direct observations and elicitation of resident perceptions of neighborhood boundaries represent alternatives to predefined boundaries but are generally dismissed as subjective, inconsistent (internally and with respect to secondary source data), and/or labor intensive.

This article describes methods used for defining neighborhood boundaries for research purposes that supplement census block group information with reviews of publicly available land-use data and systematic observation. The objective in combining these methods was the selection of neighborhood boundaries consistent with census data, study recruitment goals, and residence patterns. A secondary objective was to ensure that the range of environmental characteristics to be examined in the research was present in study neighborhoods.

**Background**

Inner-City Mental Health Study Predicting HIV/AIDS, Club and Other Drug Transitions (IMPACT) is a 5-year, multilevel neighborhood study aimed at (1) identifying associations among features of the social and physical urban environment, sexual and drug use risk behavior, and HIV prevalence in New York City (NYC) neighborhoods and (2) evaluating interrelations between features of the urban social and physical environment that shape individual sexual and drug use risk behavior and that may facilitate identification of high-risk areas that can benefit from intervention.

To ensure the inclusion of neighborhoods with a range of environmental characteristics, 36 NYC neighborhoods were included in the study: three neighborhoods in each of twelve larger communities. The 12 communities were selected for high rates of HIV infection and heroin overdose. They are primarily but not exclusively low income and are geographically disperse: four of the twelve are in the borough of Manhattan (East and Central Harlem, Chelsea, and the Lower East Side); three in the Bronx (South Bronx, Hunts Point, and Tremont); three in Queens (Long Island City, Corona, and Jamaica); and two in Brooklyn (Bedford-Stuyvesant and Bushwick) (Table 1). IMPACT was approved by the Institutional Review Board at The New York Academy of Medicine.

Study design called for recruitment of 3000 participants, approximately 83 per neighborhood, between January 2005 and December 2008. A systematic random sampling procedure was developed to approach pedestrians in each of the defined neighborhoods for screening and study enrollment. Potential participants were required to be residents of the neighborhood or to spend at least 50% of their time there. Interested individuals meeting these criteria were screened for drug use so as to be categorized as: (1) current drug user (ever injected); (2) current drug user (never injected); (3) former drug user; and (4) never used drugs (except alcohol and marijuana). Within each neighborhood, approximately one quarter of study participants were required to fall within each of these categories. Eligible individuals who agreed to participate completed a comprehensive interview probing demographics, HIV risk behaviors, social support, and

Table 1. Neighborhood characteristics from 2000 census

<table>
<thead>
<tr>
<th>Neighborhood Characteristics</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Borough</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bronx</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Brooklyn</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>Manhattan</td>
<td>12</td>
<td>33</td>
</tr>
<tr>
<td>Queens</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td><strong>Poverty rates</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤25%</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>&gt;25%, ≤50%</td>
<td>25</td>
<td>69</td>
</tr>
<tr>
<td>&gt;50%</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td><strong>Total population</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤3000</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>&gt;3000, ≤6000</td>
<td>18</td>
<td>50</td>
</tr>
<tr>
<td>&gt;6000, ≤9000</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>&gt;9000</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td><strong>Race/ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American/Black</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10%</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>&gt;10%, ≤50%</td>
<td>20</td>
<td>56</td>
</tr>
<tr>
<td>&gt;50%</td>
<td>13</td>
<td>36</td>
</tr>
<tr>
<td>Latino</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10%</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>&gt;10%, ≤50%</td>
<td>17</td>
<td>47</td>
</tr>
<tr>
<td>&gt;50%</td>
<td>17</td>
<td>47</td>
</tr>
<tr>
<td>White</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10%</td>
<td>29</td>
<td>81</td>
</tr>
<tr>
<td>&gt;10%, ≤50%</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>&gt;50%</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤10%</td>
<td>32</td>
<td>89</td>
</tr>
<tr>
<td>&gt;10%, ≤50%</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>&gt;50%</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
psychosocial stress. Study design also included blood tests of participants for HIV, hepatitis B, hepatitis C, as well as syphilis and urine testing for chlamydia, gonorrhea, and trichomonas. Participants were paid $20 for their time and effort.

Attributes of the neighborhood social and physical environment that have been associated with a variety of health-related behaviors and outcomes\textsuperscript{15,17} were gathered from the census and other secondary sources (e.g., population, income, race and ethnicity, use of public benefits) and from annual street-level observation and assessments\textsuperscript{40} (e.g., land use, community institutions, vacancy and construction of buildings, green space, and upkeep of private and public spaces).

Methods

An essential step in the implementation of the study was to delineate boundaries for each of the 36 neighborhoods that could then be used to establish sampling frames and as key units of analytic interest. Consistency regarding census data was considered necessary, so that sociodemographic and other publicly available data could be incorporated into analyses. The degree to which any particular census-defined area was appropriate for study purposes could not be determined from secondary source data alone, however, some type of check was essential. Feedback from residents was an option, but the processes required for eliciting and reconciling resident views were considered too time consuming, given the number of neighborhoods included in the study\textsuperscript{41} instead, reviews of published land-use data and systematic street-level observations, carried out between February and July 2005, were used.

The specific procedures used to delineate neighborhoods were: (1) identification of three general locations in each of the twelve broader communities within which (or around which) specific neighborhoods could be identified. Criteria used to select these more general locations focused on potential for recruitment (based on field visits), clear placement within the broadly defined community (rather than straddling a border commonly used for city planning or other purposes), and a requirement that none of them be contiguous; (2) development of maps, using specialized mapping software (ArcGIS; ESRI, Redlands CA), of the targeted and outlying areas, indicating census block group boundaries; (3) review of land-use maps and census tract data available from the NYC Department of City Planning\textsuperscript{42} to ensure, prior to going into the field, that particular census block groups were residential and likely to have sufficient population (at least 1000 per census tract) for recruitment purposes; and (4) field visits and observation in each of the targeted areas, as described in the information to follow.

Field visits to prospective neighborhoods involved observation and written documentation (primarily on the maps indicating block group boundaries) of environmental features of each street considered as a neighborhood boundary and 75% to 100% of internal streets. Similar observations and documentation were made of streets near to the projected boundary, in order to assess logical expansions of, or shifts in, neighborhoods. Observations were made on foot and required an average of 2 hours (range, 30 minutes to 4 hours) per neighborhood, with additional time required for areas that seemed appropriate according to census data but were judged inappropriate when observed (generally due to very low levels of pedestrian traffic). To ensure consistency across neighborhoods, observations were guided by a preidentified list of environmental features and conducted by a single individual, experienced in community-level and qualitative research. Input and consultation regarding the appropriate size of neighborhoods, adequate distance between neighborhoods, and acceptable levels of within neighborhood heterogeneity was provided by study investigators and other study staff.

Although size of the neighborhoods was not strictly defined in advance, each neighborhood had to be large enough so that recruitment would be feasible, yet small enough to allow for comprehensive mapping of each block in the study neighborhoods. The neighborhoods also had to be of a size so that there would be some relation to resident perceptions of the neighborhood, and environmental factors considered in the analysis were near enough to all blocks included to be relevant to survey participants.

In making the final decisions regarding neighborhood boundaries and the inclusion or exclusion of particular streets, a number of guidelines, listed in order of priority, were used.

1. Census block group boundaries. Census data have been used in the study to describe a number of sociodemographic neighborhood characteristics; consequently, block groups, the smallest unit for which census data are tabulated,\textsuperscript{43} were only divided (i.e., part of the block group was included in the study neighborhood and part excluded) if they were on the perimeter of the defined neighborhood and if the excluded portion had neither residences (and therefore no residential census data attributed to them) nor appropriate pedestrian traffic (e.g., streets with warehouses).

2. The potential for efficient recruitment of appropriate participants. Block groups without sufficient population density, lacking local pedestrian traffic, and/or outdoor places where adults congregate (such as parks) were excluded from consideration, because recruiting participants would require a disproportionate allocation of project resources.

3. Evidence of homogeneity. Relative homogeneity was considered important to the definition process, because it is a characteristic commonly used by residents to define the boundaries of their neighborhood. In addition, homogeneity facilitates the appropriate attribution of census data to study participants recruited in particular neighborhoods. Homogeneity in economic status (as evidenced by housing type and housing maintenance) was the primary consideration. Homogeneity of other characteristics (e.g., ethnicity) was a secondary consideration, evident in the majority but not all defined neighborhoods.

4. Connectivity to and/or segregation from adjacent blocks. In a further effort to define neighborhood boundaries in a way that would most likely be consistent with resident perceptions, observable physical obstructions were used as neighborhood boundaries if present (e.g., major thoroughfares, railroad tracks, and institutions covering mul-
tiple blocks). Where obvious boundaries were absent, blocks appearing representative of the surrounding area were selected.

5. Heterogeneity of social and physical characteristics across study neighborhoods. Differences in ethnicity, housing type (e.g., public housing, private subsidized, other private), land use (mixed use or purely residential), proximity to public transportation, and environmental features likely to affect quality of life (e.g., elevated trains, parks, vacant lots) were necessary for study objectives but are widespread in NYC neighborhoods. Because diversity arose naturally, it was a secondary consideration in the neighborhood definition process. Minimizing the significance of particular environmental attributes to neighborhood definition should reduce bias in analyses related to those features.

Results

Defined neighborhoods range from 1 to 8 census block groups, with populations (according to the 2000 census) 44 ranging from 2252 to 11,503 (mean=5320) (Table 1). Most neighborhoods included in the study had a majority black (12 neighborhoods) or Latino (17 neighborhoods) population. The percent of population below poverty ranged from 9.7% to 65.9% (mean=40.0%), compared with 21.2% for NYC as a whole. 45 Eighteen neighborhoods included public housing, in all (five neighborhoods) or in part (thirteen neighborhoods). Three neighborhoods were ten minutes or farther from a subway stop, eight included or were bordered by streets with light industry (e.g., warehouses, auto painting), and eleven were bordered by or incorporated an elevated train or highway.

Descriptions of four IMPACT neighborhoods are presented in the information to follow. These descriptions, provided as examples, demonstrate that the use of observation as part of the boundary definition process facilitates the identification and grouping of census block groups having attributes consistent with the concept of “neighborhood” and with study objectives.

1. South Bronx 1: The equivalent of eight city blocks and two census block groups, this neighborhood included a public housing development at one end and a busy commercial and transit hub at the other. The streets between had older residences (generally, two- or three-family homes) and small businesses targeted to local consumers. The bordering areas that were excluded from the defined neighborhood following field observations contained a main shopping area for the South Bronx (considered inefficient for recruiting neighborhood residents) and several blocks of new row houses with gated driveways and entries (considered inefficient for recruitment due to an absence of pedestrian traffic; also considered to be socioeconomically disparate from included blocks). Triangular in shape, the two lengthy borders were large commercial streets.

The third border was chosen because housing type and land use was significantly different in the adjoining census block group.

2. Hunts Point 2: Hunts Point 2 was striking in the number of environmental characteristics it contained that would be considered detrimental to health and well-being. It sat within a small, industrial South Bronx peninsula (Hunts Point), separated from other Bronx neighborhoods by an expressway. Hunts Point houses multiple waste transfer stations and a large city-owned food distribution center. There is no subway service on the peninsula, and businesses catering to local residents are sparse; residents must walk under the expressway for the subway and most shopping. The equivalent of nine city blocks and five census block groups, Hunts Point 2 partially surrounded a multiblock juvenile detention center. The other borders were major thoroughfares, including one border street that was a heavily traveled truck route to the food distribution center. The far side of this thoroughfare was excluded after review of land-use maps, due to the sparseness of housing. The neighborhood was mixed use, with significant light industry and housing. The housing, which included some new construction, was relatively well maintained. The neighborhood had a large community garden, also well maintained.

3. Central Harlem 1: A traditionally African American community, Central Harlem has been economically depressed for decades, yet by the time of this study, it was experiencing gentrification and growing economic disparities. A number of easily observed conditions in Harlem appeared to make the gentrification process unique, including (1) architecturally rich but poorly maintained (including vacant) housing, meaning that purchase prices were likely to be low and refurbishment could meet the expectations of new buyers, (2) empty lots that allowed for significant new construction; and (3) the extreme disparity in value of new (or newly refurbished) and older housing. Historically, poverty rates have been high in Central Harlem, but at the time of the study, luxury apartments and brownstone shells each sold for over $1 million. 46 The neighborhood designated as Central Harlem 1, consisting of six city blocks and three census block groups, typified the changes occurring in Harlem. Virtually every street had some combination of buildings that appeared to have been continuously occupied and modestly maintained, vacant buildings, buildings under construction, and newly built or refurbished luxury buildings. The blocks designated as Harlem 1 were not distinct from those that surround them, rather they were chosen as representative of the wider area. The observation process resulted in a shift in blocks included in the neighborhood. A large apartment complex was excluded in favor of alternate blocks that were consistent with those described previously.
Discussion

Despite the growing body of literature exploring neighborhood effects on health, little attention has been paid to processes for neighborhood delineation. Most researchers use predefined boundaries consistent with secondary source data, despite shortcomings, which include possible obfuscation of local level variability and inconsistency with respect to the social or environmental factors being examined. For research involving primary data collection, including surveys or systematic documentation of neighborhood features, secondary source boundaries pose an additional problem in that it may be impossible to discern, prior to the start of data collection, the suitability of such areas for participant recruitment and/or examining environmental features of interest.

Conclusion

Neighborhood boundary definitions are subjective, varying even among residents. Consequently, there is no one precise way to delineate a “neighborhood.” A multistep process that begins with the development of census maps and reviews of published land use as well as census data for targeted communities, and continues with systematic street-level observations, is a relatively efficient methodology that allows for consideration of a range of factors commonly used in neighborhood definition, including sociodemographic characteristics, housing type, and obstructions to pedestrian traffic. An added value of the street-level observations is the opportunity to collect qualitative information that can be used as context for quantitative analyses.

There are limitations to the method, however, including its subjectivity and lack of resident perspective. Resident perspective could be added if time and funding permit, although reconciliation of diverse perspectives is likely to represent a challenge. Subjectivity cannot be eliminated, however neighborhoods defined this way can be compared with block group combinations identified by cluster analyses of census data.

The research utilizing the defined neighborhoods is just beginning. We anticipate that manuscripts describing preliminary results will be published within 1 year; final results will be available once data collection and analysis is complete. As the study progresses, the suitability of neighborhoods defined through this process will be assessed relative to study objectives and through more detailed observation of neighborhood characteristics including, if funds are available, resident perceptions of neighborhood attributes and boundaries.

We thank Vincent Edwards, who assisted with the preliminary mapping of NYC neighborhoods and Shannon Blaney, who prepared numerous census maps and compiled census data for neighborhoods included in the study. This study is supported by the National Institute on Drug Abuse (DA017020 and DA018061) and the National Institute of Mental Health (MH08192).

No financial conflict of interest was reported by the authors of this paper.

References


39. NYC Department of City Planning. Socioeconomic profile social characteristics—New York City; 2002.


45. NYC Department of City Planning. Socioeconomic profile social characteristics—New York City; 2004.