Urban Neighborhood Information Systems: Crime Prevention and Control Applications

April Pattavina, Glenn Pierce, and Alan Saiz

Over the last fifteen years, advances in computer technology have begun to significantly influence the way criminologists and policy makers think about crime and justice. One of the more influential technological advances has been the development of computer mapping and Geographic Information Systems (GIS). The adoption of computer mapping technology by many criminal justice agencies and academics has helped to move criminal justice policy and research in a direction that places increasing emphasis on the relationship between crime and the environmental context in which it occurs.

This direction has also been supported by other important trends such as the growing body of research demonstrating a link between crime and the environment and the shift in policing philosophy to a more community- or problem-oriented approach. These forces have created an abundance of opportunities to use the latest mapping technology in powerful ways to assist those interested in addressing public safety issues in local communities.

To support community-based research, many agencies have developed some Geographic Information Systems (GIS) capacity. A basic GIS system includes individual and/or area-level data with geographic indicators along with mapping software capable of producing geographic displays of information. Some systems also include statistical tools for advanced spatial analysis. There is a
growing body of literature discussing the development and application of GIS systems for criminal justice agencies, especially police departments.

One of the major topics discussed in the latest research is the type of information that should be included in a comprehensive GIS system. Many agree that data in a GIS system should be available at several geographic levels and that it should come from a variety of sources. At the most basic geographic level is address-level data. Individual-level crime data may include police 9-1-1 calls for service, crime incidents reported by the police, criminal offenders, and police coverage. Data should also be included from other community agencies. Of primary interest are datasets that include locations of schools, parks, abandoned buildings, and other information that helps to define the physical landscape of an area.

There is no question that understanding crime in its immediate physical context is extremely useful for community problem solving. In fact, much of the discussion involving the development and use of GIS systems in criminal justice focuses on the use of address-level data collected from police and other public agencies. The issues surrounding the development of neighborhood- or area-level databases, however, have not been given much attention in the criminal justice literature.

Although many acknowledge the importance of area-level data for studying the geographic distribution of crime, few go beyond this acknowledgment to discuss either the technical aspects of developing an area-level data GIS component or the applied uses of this type of data for criminal justice. This is unfortunate since several cities have developed neighborhood indicator databases, many of which contain useful information for studying public safety issues in urban areas.

The purpose of this article is to chronicle the development of an interdisciplinary, integrated neighborhood-level database for the city of Boston (funded by The Boston Foundation) and discuss the potential applications of this database to a wide range of criminal justice problems and initiatives. Chronicling such projects is important because the strategies undertaken can be shared with those interested in pursuing similar projects. The challenges we encountered and the lessons learned may be particularly useful for cities that are currently developing community-based public safety databases. Projects currently underway include the National Institute of Justice-sponsored Strategic Approaches to Community Safety Initiative (SACSI) and the Computer Mapping, Planning and Analysis of Safety Strategies (COMPASS). These projects are being developed in several cities and emphasize a “multi-agency collaboration to data-driven problem solving” (1).
In the sections that follow, we will: 1) discuss the need for creating a Boston neighborhood database; 2) chronicle the development of the Boston Children and Families Database; 3) describe the contents of the database; and 4) discuss potential uses for neighborhood crime prevention and control strategies.

The Need for a Boston Neighborhood-Level Database

The need for neighborhood-based information for the city of Boston came to the attention of the Boston Foundation in the early 1990s. The Foundation had recently launched what was then referred to as the Persistent Poverty Project. One of the goals of the project was to prepare a status report on poverty and related conditions in Boston neighborhoods. An even more ambitious goal, however, was to compile and transfer the information collected about neighborhood conditions to community residents, organizations, agencies and researchers for the purposes of community advocacy and planning.

Collecting neighborhood-based information became increasingly central to this endeavor because the federal and local governments were demanding that to secure funding for local community-based projects, potential grantees had to exhibit a detailed quantitative understanding of community problems. Potential grantees were also being asked to demonstrate partnerships among public and private agencies in implementing proposed community-based projects and programs.

As the foundation began to look into the relationship between local community characteristics and social problems, it became apparent that information necessary to study these trends was either seriously lacking or difficult to access due to technical constraints. The data set most readily available (and, therefore, most widely used) for community or neighborhood-level research was the U.S. Census, which is conducted every ten years. Indeed, much of the important research done on communities and crime in the 1980s utilized U.S. census data to measure community characteristics.

Although the census provided information on a breadth of topics, it lacked the depth necessary to study issues particular to certain cities or neighborhoods. Moreover, the census is conducted every ten years making it difficult to monitor neighborhoods over time. Clearly, alternative, consistent and sustainable information gathering strategies were needed if communities were to gain a full understanding of the problems they confront and the ability to monitor them over time.
Development of the Neighborhood Database

Recognizing the need for neighborhood-level information along with the realization that most of those interested in community development did not have the time or technical capacity to collect and compile such information, The Boston Foundation launched an initiative to design a neighborhood-level database that would include pertinent information about Boston neighborhoods. The questions initially confronting the Boston Foundation about developing the database were: 1) what information should be included in the database? 2) where would the data come from? 3) how should neighborhoods be defined? and 4) how should the data be accessed?

What information should be included in the database?
The goal of the project under which the database was to be developed was to understand the nature of persistent poverty and disadvantage in Boston neighborhoods. It was decided that to serve this purpose, any community-based information relevant to the health and well-being of children and families living in Boston would be considered for inclusion. Meetings were held with community leaders, residents, researchers and practitioners to determine what types of information should be included in the database. This process was very useful in that it gave all those interested in doing community-based research an opportunity to voice their information needs and their concerns about how the information would be made available to them. An initial “wish list” of neighborhood indicators was created as a result of the meetings.

Where would the information for the database come from?
Representatives of public agencies were also invited to attend meetings to discuss the types of information their organizations would be willing to make available for the database. Only automated information would be considered for inclusion in the database. The agencies that ultimately contributed were Boston Public Schools, Boston Police Department, Massachusetts Office of Child Care Services, Massachusetts Department of Health, Massachusetts Division of Health Care, Finance, and Policy, Coles Business Directory, Boston Assessing Department, and the Federal Financial Institution Examination Council. (See Table 1.) Some of the agencies gave the information for free, and others charged a fee for their data services. When necessary, confidentiality and commercial agreements were also reached. In addition to the agency data, a wide range of 1990 census indicators were selected for inclusion in the database.
It is very important to emphasize that the incentive for this project was not to enhance public safety *per se*. Thus, law enforcement in the city did not have a lead role in this project. They were seen as no more important to the success of the project than other social service agencies.

The meetings held between agencies, researchers, and community representatives were beneficial to the project and to all involved. For community representatives, this process enhanced their understanding of the technical and practical issues related to the development of the database. For example, they began to realize that the inability of agencies to provide the level of detail they requested for some of the indicators was often because the data simply did not exist in any automated form. This served to ease any suspicions community representatives may have had about public agencies deliberately withholding useful information.
Agencies also became less guarded as a result of the meetings. As they began to understand that the motivation behind this project was to improve neighborhoods through problem solving, they began to trust that any information they provided would not be used for purposes of undermining or publicly criticizing the contributing agencies. This process also helped many agencies reconsider their information-gathering philosophies. Many agencies were collecting information solely for administrative or accounting purposes. Getting involved in this project helped them think about data quality issues as well as other research opportunities that would be possible for their own agency if more detailed information were to be automated.

How should neighborhoods be defined?

Prior research has shown that there is considerable variation in the way in which neighborhood boundaries are determined. Based on meetings with the database group, it was apparent that the source of this variation depended largely on who was defining the neighborhoods and for what purpose. For example, community activists defined neighborhood boundaries based upon resident perceptions of where neighborhood boundaries lie, whereas public agencies responsible for economic development and the allocation of governmental resources defined neighborhoods based upon administratively defined boundaries.

The challenge was to provide the data in the most geographically flexible format so that multiple neighborhood boundary definitions were possible. It was also necessary to make the data geographically consistent with the U.S. Census since information from the census was selected for inclusion in the database. Therefore, it was determined that, when possible, the data be made available to the public at the block group and census tract levels. Providing data at multiple levels of aggregation offered convenience to the user as well as flexibility in serving a multitude of neighborhood boundaries. Using census-based boundaries also made computer mapping easier because census-based maps were readily available.

Agency data came configured in a variety of ways. For those agencies that provided individual-level data, it was necessary to geocode the data. Geocoding is the process of assigning geographic coordinates to data elements, often using mapping software. Although some agencies may have been and continue to be geocoding data for their own research, we did not receive any address-level data that had been geocoded. Subsequently, the data would be aggregated to the census block group and census tract levels.
A few agencies, bound by agency policy and confidentiality, were only able to provide data aggregated to larger geographic areas such as census tract or ZIP code. Agency data formats also varied. Some of the agencies, often the ones with the greatest number of records, provided data in ASCII text formats. Other agency data came in various database or spreadsheet formats.

Northeastern University’s Center for Criminal Justice Policy Research was contracted to provide the technical development of the database. The initial team consisted of the director of the research center, a mainframe computing programmer, a GIS specialist, and a PC applications specialist. The team was responsible for acquiring data from each agency, documenting the contents of each agency database along with potential uses and limitations, geocoding the data, aggregating to various area levels, and integrating all the indicators into a final database. The average project costs were approximately $70,000 per year. Northeastern also provided additional institutional support.

**How will the database be accessed?**

It was decided early that the data would be made available to anyone interested in understanding the conditions facing Boston neighborhoods. The issue was to figure out how to organize the data in a way that was easy to access for the majority of potential users. We chose to compile the data in a Lotus 1-2-3 spreadsheet format since many potential users were familiar with spreadsheets.

Despite efforts to make the data as easy to access as possible, the large amount of information contained in the database still proved to be overwhelming to many who were interested in using the database but had limited technical and analytic skills. The Lotus Corporation philanthropy group generously donated software and training to community organizations and representatives who were in need of assistance. A user interface was created to automate the extraction of selected indicators and compile them in a separate customized file to be used for further analysis.

**Database Content**

The current database has over 1,000 neighborhood indicators. The database has included yearly data from most of the contributing agencies since 1990. Some agencies contribute data every other year. Since space limitations prohibit a complete list of all of the indicators, a sample selection is provided in Table 2. Some crime indicators
include number of assaults, robberies, and drug incidents. Examples of economic indicators include property values and the number of persons receiving public assistance. Health indicators include, but are not limited to, the number of low birth-weight babies and the number of deaths. Examples of community assets include the number of retail and recreational establishments. Some youth indicators include the number of days absent from school and median test scores, and family indicators include the number of single-parent families.

### TABLE 2

**Selected Indicators from the Boston Children and Families Database**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crime Indicators: Number of</strong></td>
<td></td>
</tr>
<tr>
<td>Homicides</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Aggravated assaults with a gun</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Aggravated assaults with a knife</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Simple assaults</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Armed robberies – street</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Armed robberies – commercial</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Rapes</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Drug calls</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Minor disturbances</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Burglaries – commercial</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Burglaries – homes</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td>Priority 1 calls</td>
<td>Boston Police Department</td>
</tr>
<tr>
<td><strong>Economic Indicators:</strong></td>
<td></td>
</tr>
<tr>
<td>Number receiving public assistance</td>
<td>1990 U.S. Census</td>
</tr>
<tr>
<td>Real estate values</td>
<td>Home Mortgage/Disclosure</td>
</tr>
<tr>
<td><strong>Physical Indicators:</strong></td>
<td></td>
</tr>
<tr>
<td>Number of commercial parcels</td>
<td>Boston Assessing Department</td>
</tr>
<tr>
<td>Number of residential parcels</td>
<td>Boston Assessing Department</td>
</tr>
<tr>
<td>Number of residential parcels in poor condition</td>
<td>Boston Assessing Department</td>
</tr>
<tr>
<td>Land value</td>
<td>Boston Assessing Department</td>
</tr>
<tr>
<td>Property value</td>
<td>Boston Assessing Department</td>
</tr>
<tr>
<td>Number of vacant properties</td>
<td>1990 U.S Census</td>
</tr>
<tr>
<td><strong>Health Indicators:</strong></td>
<td></td>
</tr>
<tr>
<td>Number of low birth-weight babies</td>
<td>MA Department of Health</td>
</tr>
<tr>
<td>Number of mothers who smoke</td>
<td>MA Department of Health</td>
</tr>
<tr>
<td>Number of hospital admissions</td>
<td>MA Division of Health Care, Finance, and Policy</td>
</tr>
<tr>
<td>Number of deaths</td>
<td>MA Department of Health</td>
</tr>
</tbody>
</table>
Community Asset Indicators:

- Number of retail establishments: Coles Business Directory
- Number of lodging establishments: Coles Business Directory
- Number of recreation establishments: Coles Business Directory

Youth Indicators:

- Number of public school children in special education: Boston Public Schools
- Number of days absent for public school children: Boston Public Schools
- Number of dropouts—public school children: Boston Public Schools
- Number of suspension days—public school children: Boston Public Schools
- Median standard test scores: Boston Public Schools

Family Indicators:

- Number of single-parent families: 1990 U.S. Census
- Number of public school children receiving food service: Boston Public Schools
- Group day-care capacity: MA Office of Child Care Services
- Family day-care capacity: MA Office of Child Care Services

Uses for the Database

Even though the database was not specifically designed for use by criminal justice specialists, there are many analytic possibilities for this database that have both theoretical and practical relevance for understanding the relationship between community characteristics and crime in the city of Boston. Among the more traditional research studies for which the database can be used are those that seek to determine statistical relationships between community characteristics and crime using a large number of neighborhoods. Indeed, prior quantitative research of this nature has demonstrated rather consistently that crime and other social problems tend to cluster in the same neighborhoods.

Community-level research has certainly highlighted the need to consider community characteristics in understanding crime, but translating this research into local policies and practices requires neighborhood-specific analysis. While prior research may guide the selection of indicators to consider in an analysis of criminal behavior at the neighborhood level, it is important to remember that emergent statistical patterns are not indicative of all neighborhoods. Focusing on one or several neighborhoods may help to identify relationships and shape policies specific to those neighborhoods.
Since the goal of the project under which the database was developed was to assist local community improvement efforts, it is necessary to examine how the database can support neighborhood-specific analysis. We argue that this database can be particularly useful for the police and other criminal justice professionals interested in local crime prevention and control strategies. Its importance for the police lies in the opportunities it presents for studying neighborhood-based crime patterns, conducting neighborhood-based needs assessment, and evaluating crime prevention and control programs.

Given that the data are in aggregate form, confidentiality is not a major problem as is the case with individual or address-level data. Thus, access to the data is not limited to a select few. The advantage is that anyone interested in looking at neighborhood crime patterns can use the data to pursue her own research questions. This was an important feature of the database since community empowerment was a major goal.

**Neighborhood Crime Patterns**

The police incident data in the database include over twenty categories of criminal behavior. Having crime information available at such a level of detail offers opportunities to analyze crime-specific patterns for a specific neighborhood or across many neighborhoods. Relationships among several crime types across neighborhoods can also be examined.

The database includes the total number of units aggregated to the neighborhood. So there are several options to use for denominators when calculating crime rates. For example, the assessor data provides the number of parcels for each neighborhood and the census data has the number of persons. This flexibility offers different strategies for measuring crime rates. For example, the number of residential burglaries could be divided by the number of housing units to create a rate. Similarly, the number of assaults should be divided by the number of persons to create a rate.

As previously mentioned, the database is intended to provide information to be used for community-based analysis and planning. It is not intended to provide criminal incident-level information on a real-time basis for immediate problem identification. Indeed, hotspot analysis is quite effective for that type of tactical analysis and enforcement. However, we must also pay attention to the fact that when addressing crime volume, some neighborhoods have higher crime rates than others, regardless of whether or not hot spots are present. Thus, by examining crime in terms of area volume over time, useful patterns and relationships may emerge that may be missed by
examining only address-level data. Moreover, neighborhood-level data may be used to support hot-spot analysis by providing a social context with which to examine real-time point patterns and hot-spot analysis.

**Neighborhood Needs Assessment**

Not only is it necessary to analyze crime patterns at the neighborhood level, but as we have already established, it is also important to determine what factors may be associated with local crime patterns. This type of needs-based analysis can be useful for community policing initiatives. Computer mapping proves especially useful for this purpose.

For example, a descriptive map of larcenies across neighborhoods may reveal that certain neighborhoods experience higher levels than others. Although this information alone is useful in identifying which neighborhoods have higher crime levels, the database also offers analytic capability. By layering additional information onto the property crime map, neighborhood characteristics that may occur in tandem with crime can be explored. To illustrate this capacity, the number of high school suspension days may be overlaid onto the larceny crime map. Figure 1 reveals that in the central part of the city, certain neighborhoods (measured as census tracts) around Franklin Park have high levels of larcenies and high numbers of high school student suspension days. Of course it is important that this information be interpreted in an appropriate manner. Based on this bivariate relationship alone, one would be committing logical fallacy if one interpreted the relationship to mean that children suspended from school are responsible for the larcenies in these neighborhoods. It could just as easily be interpreted to mean that children suspended from school are the victims of larceny. In this case the map simply provides a basis for asking questions about the possible nature of the relationship in those neighborhoods.

Perhaps the best sources to further explain this relationship are the police who patrol the neighborhoods. It is they who work in these neighborhoods and interact with the residents. Through their daily experiences and observations on the job, they could probably offer significant details about any relationships that may be observed involving their neighborhoods. It seems realistic to assume that they would be able to shed some light on whether suspended children are more likely to be victims or perpetrators of crime in their neighborhoods.

Given the large number of indicators available in the database, there are many possibilities for exploring relationships in a manner similar to the one just illustrated. Results could be used for police
resource allocation, program development, and partnership building. Expanding upon the school-suspension/property-crime example, the realization that several neighborhoods demonstrate the same relationship patterns may facilitate discussion about the problem and possible solutions among police and school representatives who serve similarly affected neighborhoods. Perhaps a crime prevention program developed in partnership with residents and the school department could be designed to keep children in school. By working together to address the problem, interagency cooperation and information sharing can be facilitated.

FIGURE 1
Number of Larcenies by Number of High School Suspensions, 1994
Program Planning and Evaluation

Since the database maintains the same neighborhood indicators over time, it can be a valuable and convenient resource for community-based program evaluation. There are three advantages of this type of database for program evaluation. First, since the database is longitudinal, indicators can be measured before and after program implementation. Determining before-to-after change in program outcome indicators is an important step in assessing a program’s impact.

Second, given that the data are available for all neighborhoods in the city, neighborhoods found to have similar measures on specified indicators that do not have the given program can possibly be used as matched controls. Third, given the number and variety of indicators available, tracking change in other indicators may help to identify or eliminate potential confounded relationships.

Keeping this source of information readily available can prove invaluable for many crime prevention and control programs. Often, evaluation plans are not considered until after the program has been implemented. Having access to a ready-made data source that may provide useful before-and-after measures as well as create opportunities for establishing controls can save a lot of time, effort, and frustration for program evaluators.

In addition to strengthening any causal inferences that may be made regarding the impact of a particular program, the database can be also be used to estimate undesired or unintended program outcomes. In the case of crime prevention and control programs, the data can be used to determine displacement effects possibly associated with program implementation. Although crime data at the address-level mapped over short time intervals are best for determining immediate crime displacement effects of programs, the database may be used to estimate social and economic displacement associated with increases or decreases in criminal activity resulting from targeted areas. For example, if a given program is shown to reduce crime in a given area, it is important to monitor the social and economic consequences of the change. It could be that property values increase as a result of lowered criminal activity. This change may lead to the displacement of residents. Such changes may not immediately materialize, but rather occur over longer periods of time and can be monitored using the database.

These possibilities create the potential for linking criminal justice theory and practice. As Coulton describes, indicators from these types of databases can serve as contextual measures and/or outcome measures in the study of social problems. Indicators that reflect an outcome orientation include social, health, and develop-
mental measures for the population living in local areas. Outcome measures may include but are not limited to crime measures, school dropout rates, truancy rates, and teenage pregnancy.

Indicators that reflect a contextual orientation include measures of the community structure and process that are believed to affect children and family life. Contextual measures may include poverty estimates, female-headed families with children, residential mobility, school performance, property type and condition.

Of course, contextual and outcome indicator categories are not mutually exclusive. How a particular indicator is perceived depends on the type of study and the research questions being addressed. The capacity for linking theory and practice comes from the role indicators play in developing and evaluating crime prevention and control programs. If the contextual and outcome indicators used to develop a program are the same ones used to monitor that program over time, then a clearer link between context and outcome can be established. Changes in one indicator can be compared to changes in another. Proving temporal order among variables is crucial for making causal claims.

This database may be particularly useful for comprehensive community-based crime prevention and control programs such as Operation Weed and Seed. Couched within a community policing philosophy, the weeding component of such programs focuses on localized crime problems and requires address-level crime mapping to identify specific crime problems in need of direct law enforcement intervention. The seeding component involves planning programs designed to address area social and economic factors that may contribute to local crime problems.

The seeding component requires that the social and economic structure of a neighborhood be analyzed to design services based on the population needs of the designated area. For this type of planning, the database would prove useful for both program planning and evaluation. By being able to frame place-based or hot-spot information within a social context, researchers can begin to add theoretical dimensions to this type of analysis.

**Recent Uses**

There are several specific uses for the database that are worth noting. The Boston Police Department recently used the database to examine the relationship between community characteristics and community policing. Indicators from the database were merged onto an individual public safety survey conducted by the Boston Police Department. The project, funded by the National Institute of Justice, was designed to determine the individual and community factors...
associated with interest, knowledge, and involvement in community policing in Boston.

The Boston Police Department also used the database to extract social and economic data to profile their police districts. The database proved to be particularly useful since the police department was in the process of subdividing certain districts. Other projects underway involve examining the relationship between businesses and crime. The Boston Police Department has recently increased the staff devoted to crime analysis and mapping and will continue to use the database for research.

Additionally, in certain cases, permission may be granted for the use of the individual-level data provided by participating agencies. For example, Pattavina used individual-level data to determine the impact of exposure to community-level violence on the school performance of public school children in the city of Boston. She found that the level of neighborhood violence negatively affects the standardized test scores of middle and high school children, even after controlling for other neighborhood, school, and individual factors.

Other organizations such as local neighborhood-based organizations have used the data to study public health issues. One study conducted by the Codman Square Health Organization examined the issue of prenatal smoking using the database. There have been more than one hundred requests for the data. Unfortunately, there has not been a system put in place to monitor the use of the data.

Finally, the database provides unique teaching opportunities for graduate and undergraduate students. The authors have designed graduate level courses that draw heavily upon the database for teaching technical courses such as crime analysis and mapping. Students are also encouraged to get involved with the technical development of the database and are helping to put in place a system for generating the database on a regular basis.

**Limitations**

The database is subject to a variety of limitations. Each agency, of course, has its own set of limitations regarding the accuracy of the data provided. Since each agency has a different purpose for gathering data, each will have its own specific limitations regarding uses of the data. In the following section, we discuss some of the general methodological issues that may pertain to data collected by any agency.

Methodological issues include the technical aspects of data collection, geocoding, and aggregation. Data collection problems can result in reporting bias and error. One example includes the indicator measuring the number of school dropouts. At the end of the school
year, the number of official dropouts are tabulated. There are, however, instances in which children just stop coming to school and do not officially withdraw. These are considered unofficial dropouts and are not counted in the end-of-year dropout tally. A dropout figure based on end-of-year official dropouts would, therefore, be underestimated and the degree of this type of error may differ across schools and communities in an unknown way.*

The more decentralized the data-gathering strategies, the greater the chance for this type of reporting error because as the number of data collectors increases, so does the chance of differing interpretations regarding data collection and entry guidelines. This is most problematic if data collectors are community based such as hospitals or schools because these errors would become systematic, but difficult to determine.

Missing data, coding errors, and geocoding problems present additional measurement problems. These collection problems are more likely to arise in cases where information collected is intended for descriptive purposes only and does not serve any important administrative function. Other sources of error related to data collection include inconsistencies in the definition of a variable or in the number of categories of a variable from which indicators are generated. These problems are particularly insidious to our project since we are trying to provide consistent measures over time. A final comment involves the time period covered for each of the databases. Not all databases provided information for every year since 1990. Some provided information every other year in the early stages of the project. After 1995, data collection became more consistent in terms of providing yearly data from each of the agencies.

Geocoding errors arise either from inaccurate address information provided by the contributing agency or from an outdated or otherwise inaccurate map used for the geocoding process. Boston has the additional problem of having many streets with the same name. Additional geographic information for each case, such as a zip code, is then necessary to narrow the possible matches. For each database, we were often able to geocode at least 90 percent of the cases.

The database is also limited in the level of standardization that can be offered for indicators. This is a major trade-off in providing small area measures. It would be informative to standardize many indicators by important demographics such as age, race, and gender. In some cases such as the school data, the database does offer some detailed standardization. In most cases, especially in those instances where there are only a few cases in some neighborhoods, not only are

*Coulton
*To account for this problem, a dropout adjustment is made in the fall of the preceding year.
All of these problems underscore the need to be concerned with quality assurance. However, given our limited staff, we rely on a number of users who volunteered to test each version of the database before its release. While some volunteers do some initial testing, this process is not as effective as we would like in that minor errors have gone undetected until after official releases. Errors brought to our attention are immediately corrected and are available in subsequent releases of the database.

One important substantive criticism of the database is its over-emphasis on measuring the negative aspects of Boston neighborhoods. Those voicing the criticism argue that it is equally important to include positive aspects or assets of communities in the database. In response, the database project staff are seeking to incorporate more information about the positive aspects of communities in plans for future releases of the database.

**Future Directions**

Boston is just one of many cities developing neighborhood-level databases. Several other cities, including Atlanta, Chicago, Cleveland, Denver, Oakland, and Providence, have also developed neighborhood-level databases for similar reasons and all include criminal justice measures. Acknowledging the importance of these projects and their initial success, the Urban Institute brought these cities together to form the National Neighborhood Indicators Project (NNIP).

The NNIP goal has been to provide continued operational and development support to these cities and use their combined experiences to create guidelines for other cities to follow. The project has monitored the successes of the cities, documented the type of information collected, and recruited more cities into the partnership. As of this writing, there were a dozen cities involved in the NNIP project.

The Boston Children and Families Database continues to receive support from organizations such as the Urban Institute, the Boston Foundation, and Northeastern University. The people participating in the project have changed over the last decade, and governance groups have been established to guide future development of the database. Additional agencies have agreed to participate, and more indicators are being planned for future releases of the database. The major challenges that lie ahead involve including data from more agencies, sustaining and expanding financial support for the database,
maintaining current data-sharing relationships with participating agencies, and providing quality assurance and technical support to those who wish to use the database to study neighborhood problems. Studies concerning the impact of using the data on public policy are also warranted.

It is clear that neighborhood-level data collected over time offer considerable opportunities for criminal justice research, planning, and evaluation. Such databases are being developed in many cities across the United States. As technology continues to advance and more information becomes available, so too will the breadth and depth of indicators that can be added to the database. These developments will not only provide incentives for sharing information across agencies, but they will also facilitate interagency cooperation in understanding and solving local problems.
Bibliography


