Creating a Youth Violence Data System for Boston, Massachusetts

Deborah Azrael
Harvard School of Public Health, United States of America

Renee M. Johnson
Harvard School of Public Health and Boston University School of Public Health, United States of America

Beth E. Molnar, Mary Vriniotis, Erin C. Dunn, Dustin T. Duncan and David Hemenway
Harvard School of Public Health, United States of America

The Harvard Youth Violence Prevention Center has partnered with the city of Boston, Massachusetts to create a data system to track youth violence. The system provides continuing information about fighting, bullying, dating violence and gun carrying. Data from it show that many Boston adolescents are afraid on public transportation (as compared to home, school or street) and that most students widely overestimate the number of their classmates carrying guns. The system also provides data on neighbourhood variables such as collective efficacy. The system has been a success due to the mutual respect shown between the academic and city partners, and the mutual benefits it provides to both.

Keywords: youth violence, neighbourhood, surveillance system

Youth violence is a significant public health concern in the United States. Not only are US youth at high risk of homicide, but surveys of US youth have repeatedly found that the large majority of youth report having witnessed, been victimised by or having perpetrated violence in their lifetimes (Buka, Stichick, Birdthistle, & Earls, 2001). In 2005, the principal US public health agency, the Centers for Disease Control and Prevention (CDC), responded to the problem of youth violence by funding eight 'Academic Centers for Excellence' to 'study and create lasting ways to prevent youth violence' (Centers for Disease Control and Prevention, 2007).

This article describes a comprehensive collaboration between one of the centres, the Harvard Youth Violence Prevention Center (HYVPC) and the city of Boston, Massachusetts.[1] The centrepiece of the collaboration is an ongoing, multilevel, multimethod data collection and linkage effort that brings together individual and neighbourhood level data — collected from both youth and adults — as part of a broad-based effort to increase capacity to reduce youth violence in the city. Data

from the system are beginning to make it possible for policymakers, community members and academic researchers to understand the contribution of family, school, and neighbourhood factors to youth wellbeing and to design interventions to reduce youth violence, foster positive youth development and enhance youth resiliency.

Project History

In 2003, the Harvard Youth Violence Prevention Center (HYVPC) of the Harvard School of Public Health was approached by the city of Boston to help revise their Boston Youth Survey (BYS), a survey that had been given to a sample of youth in summer school and summer jobs programs intermittently since the mid-1990s. The HYVPC redesigned the survey and, in collaboration with the city, in spring 2004, administered it to a random sample of about 1,000 high school students.

The 2004 collaboration worked well and in 2005, in partnership with the city, HYVPC sought and received funding from the Centers for Disease Control and Prevention for a 5-year multitiered, multimethod research project, the Boston Data Project, designed to build capacity to reduce youth violence in Boston. The Boston Data Project (BDP) consists of two major data collection efforts: the biennial Boston Youth Survey (BYS), an in-school paper and pencil survey of Boston public high school students, and the Boston Neighborhood Survey (BNS), a random-digit-dial telephone survey of Boston adults. A key feature of data from these two surveys is that we are able to identify the US Census tract of residence for each respondent. As a result, data from the surveys can be aggregated geographically and linked to other geo-coded data sources. Assembling and linking area-level data from various Boston datasets, such as US Census and Boston Police Department incident files, is the third major component of the BDP. The three components of the system are described in more detail below.

The Boston Youth Survey (BYS)

The BYS is a biennial survey of high school students in Boston Public Schools. In 2004 and 2006, the sample was selected using a two-stage process using the sampling programs PCSchool and PCSample (Centers for Disease Control and Prevention, 2000). In Stage 1, a random sample of Boston's public high schools was selected and in Stage 2, a random sample of classrooms, stratified by grade, was selected within those schools.

In 2008, the research team sought to administer the BYS in all eligible Boston Public School high schools (n = 31). Ineligible schools were those that serve adults (i.e., night school), short-term schools (i.e., for students transitioning back to school after incarceration, or schools that house suspended students), and those that serve severely disabled youth. Twenty-two of the eligible schools participated in the survey (71%). The final sample of schools was largely representative of all schools in terms of the race/ethnicity of the students, school drop-out rates and other variables. All the schools tried to be cooperative; the primary reason for school nonparticipation was scheduling difficulties (e.g., lack of time due to the scheduling demands of standardised testing).

The survey uses passive consent procedures (i.e., students' parents were required to return a form if they did not want their child to participate) and students may

decline to participate at any time before or during the survey on the day of adminis-tration. The survey takes approximately 40 minutes to complete. The survey is administered by trained youth workers and others (e.g., researchers, graduate students, city workers) during regular class time.

In 2008, about one-quarter of students in the sampled classrooms were not in attendance on the day of the survey. This is consistent with the average daily student attendance at Boston public schools. Of those students present, about 7% refused to take the survey or were not permitted by a parent to take it (less than 1% of students). The student response rate (number of eligible surveys divided by number of enrolled students in sampled classrooms) was 69%.

The BYS provides information on: (1) indicators of positive youth development and resiliency, including school performance, extracurricular activities and contact with supportive adults; (2) risk factors for and experiences of violence, including exposure to violence and violence perpetration; (3) perceptions of neighbourhood characteristics, including collective efficacy; (4) use of community resources and (5) health behaviours (e.g., physical activity) and risk behaviours (e.g., alcohol and tobacco use).

The Boston Neighborhood Survey (BNS)

The BNS is a biennial, random-digit-dial telephone survey of approximately 1,700 adults in Boston conducted by a survey research firm. First administered in 2006, the main purpose of the BNS is to supplement BYS data with contextual informa-tion about neighbourhood-level conditions and social processes as perceived by adult residents.

Adult residents (>18 years) throughout the city of Boston are randomly selected for inclusion in the survey from a list-assisted sampling frame, with separate random probability samples (proportional to population size) selected for each of Boston's 16 neighbourhoods. As is increasingly the case for telephone surveys (Link & Kresnow, 2006), refusal rates for the BNS are relatively high, with approximately two refusals for every completed survey. Break-offs once a respondent has agreed to participate are low (< 10%), as respondents can usually complete the survey in less than 15 minutes. A verbal informed consent procedure is conducted prior to beginning each interview.

Neighborhood Data

Sources of data include the US Census, the Boston Police Department and the Boston Public Health Commission. For example, the Boston Police Department provides incident data (e.g., homicide and aggravated assault) and the Boston Public Health Commission provides emergency department and hospital discharge data.

Findings

Almost from its inception, the data system has been widely used by the city for policy planning. A few findings from the 2006 BYS that were of particular interest to the city are highlighted below.

Exposure to Violence

Understanding more about in-school youth's experience of violence is an important priority for the city. The experiences of such youth are often overlooked because both public and policy attention tends to focus on youth who are criminally involved. Each year the BYS includes questions about witnessing violence (inflicted on someone else) and about whether the respondent was personally assaulted.

In 2006, 20% of students reported having been physically assaulted (i.e., punched, kicked, choked or beaten up) within the past year. Many had been threatened or attacked with a firearm (9%), or with a weapon other than a firearm (9%). White students were more likely to have been physically assaulted (29%) than Blacks (18%) or Hispanics (15%). Over 65% of respondents reported witnessing one or more acts of violence in the past year; almost a quarter (24%) of students reported having seen someone attacked with a weapon other than a gun, and 18% reported having seen someone shot or shot at.

Feelings of Safety

Not surprisingly, given these findings, many students reported feeling unsafe in various locations in the city, though the percentage of youth reporting feeling unsafe declined markedly between 2004 and 2006. This decline is encouraging, and may be associated with overall citywide efforts to improve neighbourhood services. However, with only two years of data, it is not possible to draw strong conclusions.

As seen in Figure 1, the most 'dangerous' location from youths' perspective was the city's public transportation system, which most students take to and from school. Spurred in part by release of these findings, the Boston Public Schools made reducing youth violence in this context a priority. Any incidents that occur during
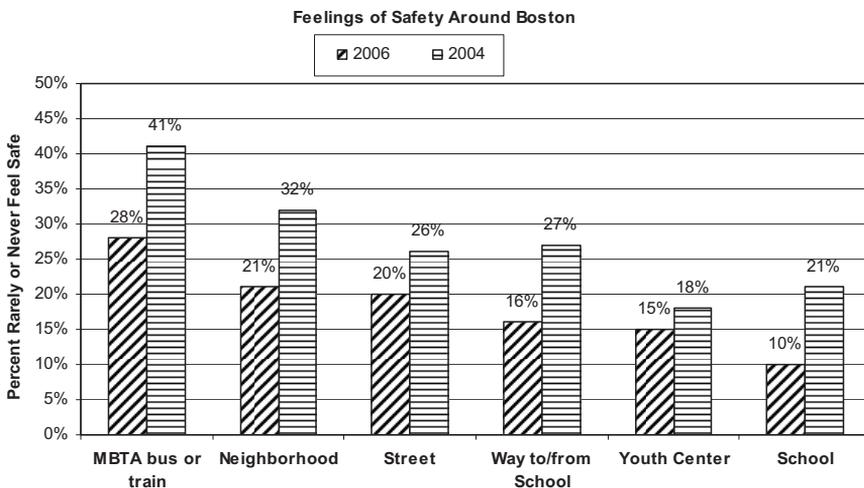
FIGURE 1

Feelings of safety.

travel to and from school are now considered within the disciplinary jurisdiction of the Boston Public Schools.

Violence Perpetration

While the Boston Police Department maintains high quality criminal justice-related data on youth violence, with the exception of limited data from the national Youth Risk Behavior Survey (a large biennial survey of high school youth conducted by the Centers for Disease Control and Prevention since the early 1990s), historically the city has had little information on youth violence falling below the threshold of clear criminality. The Boston Youth Survey provides comprehensive data on violence perpetration by students. In 2006, half of students reported minor aggression in the past month and one out of ten reported that they had attacked or threatened to attack someone (Table 1). A striking finding is that many more students report having assaulted someone than report having been assaulted themselves. A possible explanation is that students may be excluding peer and family violence from their reports of victimisation. To explore this possibility, the 2008 BYS asked students to report family, peer and other violence separately.

TABLE 1

Prevalence of Aggressive Behaviour in the Past Month, by Selected Characteristics (*n* = 1,215)

	Minor aggression: Pushed, shoved, or slapped someone	Moderate physical aggression: Hit, punched, kicked, or choked someone	Severe aggression: Attacked, or threatened to attack, someone
Total	49%	37%	10%
Age			
≤ 16 years	53%	39%	11%
> 16 years	44%*	34%	10%
Sex			
Male	47%	39%	13%
Female	50%	36%	8%*
Race/ethnicity			
Hispanic	47%	35%	11%
Black (non-Hispanic)	52%	41%	11%
White (non-Hispanic)	50%	34%	8%
Other	36%*	32%	6%
Nativity			
US-born	53%	40%	11%
Foreign-born	38%*	32%*	7%*

Note: * = *p* < .05 as assessed using chi-square tests of association; 'other' includes bi- and multi-racial students, Asians, Native Americans, and students who were neither Hispanic/Latino nor were able to classify themselves into a race category; multiple comparisons analysis showed that the 'other' category was significantly different from the other three groups. There were no other pairwise significant differences.

The prevalence of aggression varied by respondents' demographic characteristics, and most strikingly by their nativity. Students who were born outside of the United States were significantly less likely than their US-born peers to report having perpetrated any kind of physical aggression.[2]

Nearly one-third (31%) of Boston youth reported having carried a knife in the past 12 months, and 6% reported having carried a gun over that same time period. Males were significantly more likely than females to carry a knife (43% vs. 23%), or a gun (10% vs. 3%) and, consistent with earlier findings, US-born youth were more likely than immigrants to carry weapons. Even among those students who had not carried a gun in the past year, 56% of boys and 36% of girls said that acquiring a gun would be fairly or very easy. However, the large majority of youth reported that they preferred to live in a world where it was hard to get a gun (78%), a finding that held for both those who reported having carried a gun and those who had not.

Gun Carrying

Students seemed to overestimate the incidence of gun-carrying by their classmates. While approximately 5% of students reported carrying a gun in the past year, youths' median estimate of gun-carrying among their classmates was 20% (Table 2).[3] This misperception may increase gun-carrying among youth overall, as a major reason cited by youth for carrying is protection — often against peers who are perceived to be carrying guns.

'Snitching'

To counter youth social norms that stigmatise snitching, the Boston Mayor's Office has been focusing on increasing the number of youth willing to report criminal activity. In the 2006 BYS, while the majority of youth reported that they would tell some adult if they saw a crime, 28% said they would 'not tell anyone'. Those who

TABLE 2

Gun-carrying (Boston Public High School Students)

	% carry gun	'overestimate' of % of R's classmates who carry guns
Gender		
Male	10***	22~
Female	3	24
Race		
White, non-Hispanic	4	14
Black, non-Hispanic	7	24***
Hispanic	6	25***
Asian	0	8
Nativity		
US-born	7	23***
Immigrant	3~	16

Note: ~ $p < .1$, ** $p < .01$, *** $p < .001$.

reported having strong family ties and were active in their schools and communities were more likely to say they would disclose having seen a crime. On the other hand, students who were Black/African American, who had been involved with the juvenile justice system, who reported feeling disrespected by the police and who had had personal experiences with violence were the least likely to disclose having witnessed a crime.

Collective Efficacy

A crucial component of the ongoing Boston Data Project is its emphasis on the relationship between neighbourhood characteristics and youth violence, and particularly the effect of neighbourhood collective efficacy on violence.

Academic Rationale

Social scientists have sought to understand the impact of environments and health dating back to 17th century Britain when numbers and causes of death were compiled for different areas and it was observed that greater mortality occurred in towns than in the countryside (Macintyre & Ellaway, 2003). A pre-eminent contributor to early work on the social distribution of health and life expectancy was Edwin Chadwick, who showed variations among social categories and among areas in his 1842 publication *Sanitary Conditions of the Labouring Poor in Great Britain* (Macintyre & Ellaway, 2003; Chadwick, 1842). Investigations of variations in morbidity, mortality and in behaviours have continued since, igniting a debate over whether these differences represent the composition of the population living in each area, or contextual effects of the places themselves (Macintyre & Ellaway, 2003).

Focusing on contextual effects in the early twentieth century, the field of sociology began to address differences between neighbourhoods (Sampson, 2003a). For example, researchers in the 1940s researchers hypothesised that neighbourhoods with high levels of social disorganisation were most likely to experience illness and crime (Shaw & McKay, 1942). Typically, structural characteristics of communities — such as poverty, unemployment rates and the percentage of families on public assistance — have been studied for their deleterious effects. It has been relatively recent that community-level collective processes that may promote wellbeing have also been examined (Mayer & Jencks, 1989, Sampson, 2003). Sociologists such as Bronfenbrenner have placed the study of neighbourhood effects on health and wellbeing into a broader theoretical framework (Bronfenbrenner, Moen, & Garbarino, 1984; Elliott et al., 1996; Garbarino & Carry, 1997; Sampson, 1999, 2001, 2003; Sampson & Wilson, 1995) that seeks to account for the multicontextual nature of people's lives — lives influenced not only by individual and family characteristics, but also by the nature of their various social communities, including residential neighbourhoods.

A limitation of much work on neighbourhoods and community health is the paucity of relevant neighbourhood data with which to explore associations between neighbourhood characteristics and resident wellbeing. As a result, most studies of neighbourhood effects on mental and physical health have been limited to using administrative data that are relatively inexpensive to gather, but that are not designed to capture neighbourhood-level social processes and dynamics

(Raudenbush, 2003b). Over the last decade, several research projects, including the groundbreaking Project on Human Development in Chicago Neighborhoods (Earls & Buka, 1997), have created and refined measures of several of these social processes (e.g., collective efficacy, described more fully below) and developed the measurement and analytic tools necessary to explore the association between these processes and neighbourhood wellbeing.

Measures

Both the BYS and the BNS collect numerous measures of respondents' assessment of the social fabric of their communities, including collective efficacy, neighbourhood disorder and others.[4]

Collective efficacy is defined as expectations for the informal social control and mutual support of children (Sampson, 1999), a definition that emphasises neighbourhood adults' sense of active engagement toward a specific, collective goal (Bandura, 1995). Collective efficacy has been shown to be highly associated with lower levels of youth violence (Bandura, 1995; Morenoff & Sampson, 1997; Sampson, 1997, 2003; Sampson, Morenoff, & Gannon-Rowley, 2002; Sampson, Raudenbush, & Earls, 1997) including lower rates of gun-carrying (Molnar, Miller, Azrael, & Buka, 2004). Collective efficacy is also associated with later onset of first sexual activity (Browning, Leventhal, & Brooks-Gunn, 2004), better mental health (Xue, Leventhal, Brooks-Gunn, & Earls, 2005), higher standardised math tests (Emory, Caughy, Harris, & Franzini, 2008) and better school outcomes (Coley, Morris, & Hernandez, 2004). One of collective efficacy's constituent parts, informal social control, has emerged as a key mechanism in studies of the association between neighbourhood disadvantage and violence. Specifically, in neighbourhoods where people collectively watch out for misbehaviour, there are lower rates of criminal behaviour, whether or not the neighbourhood is impoverished (Sampson et al., 2002).

To measure collective efficacy, we collect data on two interrelated social processes: (a) informal social control (e.g., do adults intervene if youth are truant, loitering, or spray-painting graffiti; would neighbours work together to combat the closing of a neighbourhood service such as a fire station?) and (b) social cohesion (Sampson, 1999; Sampson et al., 2002) (e.g., how much do neighbours trust each other and share the same values?)

Defining 'Neighbourhood'

Numerous studies have explored the influence of neighbourhood characteristics, including collective efficacy, on the wellbeing of residents (Kawachi & Berkman, 2003). Within these studies, 'neighbourhood' is defined variously (Diez Roux, 2001; Messer & Kaufman, 2006; Osypuk & Sandro, 2007), and the reasons for choosing one neighbourhood definition over another are often not explicit (Osypuk & Sandro, 2007). Generally, however, neighbourhoods are defined as one or another aggregation of US census tracts (Diez Roux, 2001; Messer & Kaufman, 2006; O'Campo, 2003) — small, relatively permanent statistical subdivisions of US counties. The question of the extent to which these divisions are socially relevant, and if so to whom, is central to any enterprise that seeks to

understand the relationship between neighbourhoods *qua* neighbourhoods and the lives of the people who live in them.

Recently, several studies have assessed the degree to which individual-level characteristics predict how people (generally survey respondents) define the physical extent of their neighbourhoods. These studies have found that factors such as urbanicity, race, nativity, density of social ties and even body mass index influence the size and scope of what people call their neighbourhood (Colabianchi et al., 2007; Guest & Lee, 1984; Haney & Knowles, 1978; Logan & Collver, 1983). The Los Angeles Family and Neighborhood Study (LA FANS), for example, found that there was considerable heterogeneity in how respondents to their survey defined their neighbourhoods and that respondents' definitions were dependent on factors such as whether they worked and shopped in close proximity to where they lived, whether or not they had children and their socioeconomic status (Sastry, Pebley, & Zonta, 2002).

Other studies suggest that census tracts may not be a good proxy for neighbourhoods as experienced by residents. For example, Coulton and others (2001) compared neighbourhood maps drawn by residents to maps that used census tracts or aggregations of census tracts to approximate neighbourhoods. They found that individual maps varied significantly and that most residents' maps incorporated portions of multiple census tracts.

Despite these limitations, a number of sensible approaches to defining neighbourhood are possible and several recent studies have developed novel ways of defining neighbourhood boundaries (Chappell, Funk, & Allan, 2006; Clapp & Wang, 2006; Tatalovich, Wilson, Milam, Jerrett, & McConnell, 2006; Weiss, Ompad, Galea, & Vlahov, 2007). For example, Weiss and colleagues (2007) used a multistep neighbourhood definition process, including development of census block group maps, review of land use and field visits and observation in each of the targeted communities; Mullan and others (2004) used a technique called geographic retrofitting to define a community-oriented primary care practice's neighbourhood based on current patterns of health care by the community; and Clapp and Wang (2006) used a Classification and Regression Trees (CART) algorithm to place neighbourhood boundaries in West Hartford, Connecticut.

Neighbourhood Data From the BDP

To geo-code information from the Boston Youth Survey about the neighbourhood in which students live, but not compromise confidentiality, students are asked to provide the name of the street they live on and the nearest cross-street. Combined with information about the students' neighbourhood and postal code, this information is sufficient to assign 90% of students to a neighbourhood cluster (see following section). The Boston Neighborhood Survey of adults asks many questions about the respondent's neighbourhood, including street and nearest cross-street. The lists of phone numbers used by the firm for the survey also include census-level geographic identifiers. Neighbourhood data[5] provided by the census, and the police and public health departments are geo-coded to several different levels of specificity (e.g., census block, postal code, police district).

Defining 'Neighbourhood' in Boston

The city of Boston has a population of approximately 600,000, commonly divided into 16 named 'neighbourhoods', (e.g., Roxbury, Mattapan, Dorchester, South Boston, Back Bay and Charlestown) which vary considerably in size (from approximately 25,000 to over 90,000). Approximately one third of the city's population and almost half its youth under 18 years of age live in three of these neighbourhoods. The US Postal Service designates 31 postal codes within the city, while the US Census Bureau divides Boston into 157 census tracts (each with a population of approximately 4,000), each of which, in turn, is comprised of smaller census blocks.

From this complicated set of overlaying geographies, we have begun to identify socially relevant 'neighbourhood clusters' for the city, using a multistage, multimethod process that is in its early stages. To date, we have conducted a set of semistructured qualitative interviews with a snowball sample of key informants in each of Boston's sixteen neighbourhoods. Informants were asked to delimit their neighbourhoods by drawing a continuous line around them, using streets and natural boundaries as borders. To aid this process, respondents were shown three types of map (one from Google Earth™ and two, with different levels of detail, generated using ArcGIS™ (ESRI, 2008)).[6]

These drawings were then transferred into ArcGIS™, creating the first iteration of the first neighbourhood cluster layer. Simultaneously, several informants with broad experience of Boston policing and youth outreach assessed the maps, drawing their own boundaries. We are in the process of reconciling the boundaries of the neighbourhood clusters identified through these two processes.

Assessment

Two independent measures of the social validity of the neighbourhood clusters we have identified are (1) whether the areas are homogeneous with respect to various demographic factors (e.g., race/ethnicity, socioeconomic status) and (2) whether maps of neighbourhood clusters have face validity for the people who live in them. The homogeneity of the identified clusters is currently being assessed by calculating intercluster correlations for key racial, ethnic and socioeconomic variables. In the coming months, we will present neighbourhood cluster maps to neighbourhood groups and others to solicit their feedback. We will integrate feedback from these meetings with findings from other assessment techniques, including surveys of residents asking them to rate the appropriateness of the name and boundaries of the neighbourhood cluster in which they live on a 5-point Likert scale. Multivariate and multilevel exploration of neighbourhood characteristics and outcomes for youth at the neighbourhood cluster (n = 90) level is planned for late 2009.

Creating a Successful System

The Boston Data Project remains a work in progress, with much of its greatest potential to be realised when we begin to do neighbourhood-level analyses. Perhaps most important, as the project continues, multiple years of data will allow us to explore trends in key youth violence-related outcomes and to evaluate programs and interventions implemented with the objective of increasing neighbourhood-

level collective efficacy or reducing youth violence. Even in the interim, the Boston Data Project has already become an important source for data on youth violence in the city: the Harvard Youth Violence Prevention Center responds at least weekly to requests from policymakers, journalists, community groups and other researchers for information about and from the Boston Data Project.

Much of the success of the project is due to the city's leaders — including a mayor, police commissioner and public health commissioner — who are committed to using data in the formulation and implementation of public policy and who have participated in large-scale data collection efforts in the past. For example, for the past 10 years the city has partnered with the Boston Foundation on the a Boston Indicators Project, a 10-year-old initiative (ongoing to 2040) to track progress across numerous domains (e.g., educational attainment, health outcomes).

The city has also been actively seeking opportunities to create genuine academic–government partnerships to take advantage of the large number of academic institutions in Boston. The partnership with HYVPC has proved a good model and one that the city is potentially interested in replicating. Finally, youth homicide in Boston has increased substantially over the past few years after a long period of decline. In late 2007, the city unveiled a new public health-based approach to violence prevention, creating an Office of Violence Prevention within the Public Health Commission and making youth violence prevention a top priority. Data for policy planning and evaluation are vital to these endeavours.

The fundamental measure of the value of the Boston Data Project will be the degree to which it is successful in promoting and supporting efforts to translate research findings into policy and practice. A preliminary indication that the BDP is serving this function are the many ways we have become integrated into diverse policy development efforts across the city. For example: (1) we are advising a public–private partnership on evaluation design for a project intended to improve the ability of the city's street workers to respond to youth violence in real time; (2) we serve on the advisory and survey development committees of a city project to survey 400–600 'disconnected' youth in a problem neighbourhood; (3) we are working with a consortium of funders to develop a set of indicators to track youth violence-related issues by neighbourhood over time; (4) our data are used in presentations by the Boston Public Health Commission, Boston Police, neighbourhood health centres and many local groups. Our work has been noted outside of Boston as well; we have been asked to share what we have learned about municipal/academic partnerships in California and elsewhere.

While it is too soon to know the outcome of such activities, it is clear that the Boston Data Project and the overall model of the Harvard Youth Violence Prevention Center have become a key resource in the city and beyond.

Challenges

Our efforts to create an ongoing data system have met with numerous challenges. Initially, it was important to get the attention of the mayor and to negotiate the politics of the city. Fortunately, we were able to secure the services of a consultant who had worked for many years within city government who helped us negotiate the bureaucracy and to keep the project on the mayor's agenda.

Early in the project, it proved difficult to identify the most appropriate point of contact with the city. The Boston Youth Survey had initially been located within the mayor's office, but was transferred to the Boston Public Health Commission (BPHC) as part of a new emphasis on youth violence prevention within that branch. Among city agencies the BPHC is probably the one most familiar with surveillance systems and with data analysis and presentation.

The Boston Data Project brings together data from multiple city agencies, including the school department, the Public Health Commission, the city's Redevelopment Authority and the police department. Even with the 'blessing' of the mayor, and introductions from the Director of the Office of Human Services and later the mayor's Chief of Staff, each of these agencies had to be approached separately and 'sold' on the project. For example, our relationship with the Boston Police Department was enhanced when we were able to provide them data on what after school programs hard-to-reach youth preferred (e.g., swimming, college entrance exam prep) in each Police District to inform their summer planning. A useful mechanism for reaching multiple city departments at once was to make presentations at cabinet-level meetings.

Another challenge was to win over city staff, many of whom perceived academic researchers as being more interested in using the city to gain access to research subjects than in genuine partnership. From the beginning, we focused on putting information from the project to good use for the city, providing data and preparing fact sheets and reports with quick turnaround, making presentations when requested and assuring that the surveys include questions that are directly responsive to city concerns.

We have found that it is crucial that key city players are informed well before any findings are made public. For example, our 2006 survey showed that youth feel particularly unsafe on public transit. Before the results were released, we notified the head of the transportation system about this finding so that he could be prepared in the event of media interest and possible criticism.

Assuring that the mayor's office is up-to-date on findings from the system is particularly vital and sometimes difficult. For example, in 2005 when data from the 2004 survey were released, the Boston papers each ran prominent stories on the percentage of Boston youth who reported high levels of exposure to violence (Harvard Youth Violence Prevention Center, 2004). Because no other US cities have surveillance systems that track such youth violence-related outcomes, it was difficult, but important, to contextualise these data for the press.

Over time we have worked out an information management policy with the city; we provide our city contact with information about all upcoming presentations of data and inquiries from the press as soon as possible. We provide the mayor's Chief of Staff with email and copies of these notifications and highlight any possible areas of concern to assure that the mayor is not taken by surprise. A memorandum of agreement with the city assures that we do not report out school or neighbourhood level data without full approval from all city stakeholders.

We also receive health and public safety data from the city. Because the data belong to the agencies in which they originate, the process is complicated. For geocoded police data and with the police department's express permission, for example, we receive the data through a university-based collaborator who works with the

department. For emergency department and hospital discharge data, the Boston Public Health Commission accesses the data through an agreement with the state agency that collects them. As a result, we are required to formally request that analyses of these data be conducted by the Boston Public Health Commission.

Finally, we face various methodological challenges. For example, historical methods of reporting crime and health data pose a challenge to the choice of our Boston neighbourhoods. Police districts in the city cross neighbourhood boundaries and emergency department and hospital discharge data are readily available only by postal code. Another problem is the decreasing coverage provided by telephone surveys, due to such factors as call screening and mobile phone use. In future iterations of the Boston Neighborhood Survey we expect to move to mixed mode surveys, adding mailing and web-based components to our current phone survey methodology. Ideally, we would also broaden our sample of youth to include younger (middle school-aged) students; students attending private, parochial and charter schools; and out of school youth (including those in juvenile detention facilities).

Conclusion

The collaboration between the city of Boston and the Harvard Youth Violence Prevention Center (the Boston Data Project) is well-established and valued by the city. The data system provides a model for municipal–academic partnership and has been successful, primarily because we have been able to use it to provide actionable data to the city and other stakeholders. We have been particularly sensitive to assuring quick turnaround of data when requested, as described earlier. While many of the same questions are asked in each survey (to provide comparable information over time), the system is also flexible enough to add questions about particular city interests that arise, as they did for snitching and youth perceptions of their interactions with the police.

The project is also making a substantive contribution to science. For example, data from the project have been used to explore such issues as the relationship between hours of sleep and aggression, nativity and self-injury; discrimination and mental health outcomes among sexual minority youth; and perceptions of safety and body mass index. In the next years, we will be able to do more work using the longitudinal aspects of the data system, to help evaluate new city initiatives. We also will be able to carefully examine the relationship between neighbourhood level characteristics (e.g., collective efficacy) and violence. This latter analysis will be possible once we have completed our identification of sensible subneighbourhoods.

The Boston Data Project has already achieved a number of useful goals. First, it has demonstrated a reasonable model for multilevel surveillance of outcomes for youth that balances cost and flexibility against the complexities of real-time surveillance. Second, it has fostered numerous productive collaborations not only between the Harvard School of Public Health and the city of Boston, but across schools and departments within Harvard, and across departments within the city.

Finally, the project has the potential to enable the city, researchers and practitioners committed to reducing youth violence to explore the impact of youth violence not just on high risk youth but all youth in the city. At the time of this writing, the city is gearing up to implement a major youth violence prevention

initiative targeted at high-risk youth. The BDP will allow us to assess whether the project has impact not just on the criminally involved youth it explicitly targets, but on the overwhelming majority of noncriminally involved youth in the city and the neighbourhoods in which they live.

Acknowledgments

Endnotes

1   The HYVPC also works collaboratively with 11 grassroots community partners who have participated in the initial and ongoing development of the data system, continue to inform the direction of the HYVPC's academic research and make use of the findings from the data system for program planning and evaluation. This article focuses on the HYVPC–City collaboration.

2   Our finding that foreign-born students are less likely to report aggressive behaviour varies by nationality/ethnicity. Specifically, while Black (not Cape Verdean, non-Jamaican), Hispanic, Haitian, and Dominican students who are immigrants are significantly less likely to report aggression than their native-born counterparts, immigrant and US-born Cape Verdean, Jamaican and Asian students are equally likely to report such aggression.

3   'In the past 12 months, have you carried a gun?' (yes or no); 'What per cent of students in your school do you think have carried a gun in the past 12 months?'

4   These measures, in large part, are drawn from the Project on Human Development in Chicago Neighborhoods.

5   Accessing census data at the block, block group and tract level is difficult and time-consuming, even using the US Census web-based query system. GeoLytics makes accessing data extremely simple and also provides commonly used area-level measures such as concentrated disadvantage, the Gini coefficient and others.

6   The interview process was extremely informative and lends credence to the idea that neighbourhood clusters are difficult to define. For example, interviewees had many different, yet strikingly similar, conceptualisations of neighbourhoods. Neighbourhoods were thought of in terms of neighbourhood/civic associations, public transportation stops, churches, public housing communities, major thoroughfares and such dimensions as income and racial/ethnicity. Many respondents noted that 'neighbourhood' meant different things to different people.

References

Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 1–45). Cambridge, UK: Cambridge University Press.

Bronfenbrenner, U., Moen, P., & Garbarino, J. (1984). Child, family, and community. In R. Parke (Ed.), *Review of child development research* (Vol. 7, pp. 283–328). Chicaco: University of Chicago Press.

Browning, C.R., Leventhal, T., & Brooks-Gunn, J. (2004). Neighborhood context and racial differences in early adolescent sexual activity. *Demography, 41*(4), 697–720.

Buka, S.L., Stichick, T.L., Birdthistle, I., & Earls, F.J. (2001). Youth exposure to violence: Prevalence, risks, and consequences. *American Journal of Orthopsychiatry, 71*(3), 298–310.

Centers for Disease Control and Prevention. (2000). *PC sample/PC school: Survey TA sampling software*. Atlanta, GA: Centers for Disease Control and Prevention.

Centers for Disease Control and Prevention. (2007). *National Academic Centers of Excellence on Youth Violence Prevention*. Retrieved January 13, 2009, from http://www.cdc.gov/ncipc/res-opps/ace/ace.htm

Chadwick, E. (1842). *Report of an Enquiry into the Sanitary Conditions of the Labouring Population of Great Britain*. London: Poor Law Commission.

Chappell, N.L., Funk, L.M., & Allan, D. (2006). Defining community boundaries in health promotion research. *American Journal of Health Promotion, 21*(2), 119–126.

Clapp, J.M., & Wang, Y.Z. (2006). Defining neighborhood boundaries: Are census tracts obsolete? *Journal of Urban Economics, 59*(2), 259–284.

Colabianchi, N., Dowda, M., Pfeiffer, K.A., Porter, D.E., Almeida, M., & Pate, R.R. (2007). Towards an understanding of salient neighborhood boundaries: Adolescent reports of an easy walking distance and convenient driving distance. *International Journal of Behavioral Nutrition and Physical Activity, 4*.

Coley, R.L., Morris, J.E., & Hernandez, D. (2004). Out-of-school care and problem behavior trajectories among low-income adolescents: Individual, family, and neighborhood characteristics as added risks. *Child Development, 75*(3), 948–965.

Coulton, C.J., Korbin, J., Chan, T., & Su, M. (2001). Mapping residents' perceptions of neighborhood boundaries: A methodological note. *American Journal of Community Psychology, 29*(2), 371–383.

Diez Roux A.V (2001). Investigating neighborhood and area effects on health. *American Journal of Public Health, 91*(11), 1783–1789.

Earls, F., & Buka, S.L. (1997). *Project on Human Development in Chicago Neighborhoods: Technical Report*. Rockville, MD: National Institute of Justice.

Elliott, D.S., Wilson, W.J., Huizinga, D., Sampson, R.J., Elliott, A., & Rankin, B. (1996). The effects of neighborhood disadvantage on adolescent development. *Journal of Research in Crime and Delinquency, 33*(4), 389–426.

Emory, R., Caughy, M., Harris, I.R., & Franzini, L. (2008). Neighborhood social processes and academic achievement in elementary school. *Journal of Community Psychology, 36*(7), 885–898.

ESRI. (2008). *ArcGIS: A complete integrated system*. Retrieved November, 2008, from http://esri.com/software/arcgis/

Garbarino, J., & Carry, F. (1997). The community context of child abuse and neglect. In J. Garbarino and J. Eckenrode (Eds.), *Understanding abusive families: An ecological approach to theory and practice* (pp. 56–85). San Francisco: Jossey-Bass.

Guest, A.M., & Lee, B.A. (1984). How urbanites define their neighborhoods. *Population and Environment, 7*(1), 32–56.

Haney, W.G., & Knowles, E.S. (1978). Perception of neighborhoods by city and suburban residents. *Human Ecology, 6*(2), 201–214.

Harvard Youth Violence Prevention Center. (2004). *2004 Boston Youth Survey*. Boston. MA: Harvard School of Public Health.

Kawachi, I., & Berkman, L. (2003). *Neighborhoods and health*: Place of publication: Oxford University Press.

Link, M.W., & Kresnow, M. (2006). The future of random-digit-dial surveys for injury prevention and violence research. *American Journal of Preventive Medicine*, *31*(5).

Logan, J.R., & Collver, O.A. (1983). Residents' perceptions of suburban community differences. *American Sociological Review, 48*(3), 428–433.

Macintyre, S., & Ellaway, A. (2003). Neighborhoods and health: An overview. In I. Kawachi & L.F. Berkman (Eds.), *Neighborhoods and health* (pp. 20–42). Oxford University Press, New York.

Mayer, S.E., & Jencks, C. (1989). Growing up in poor neighborhoods – how much does it matter? *Science, 243*(4897), 1441–1445.

Messer, L.C., & Kaufman, J.S. (2006). Using census data to approximate neighborhood effects. In J.M. Oakes & J.S. Kaufman (Eds.), *Methods in social epidemiology* (pp. 209–236). San Francisco, CA: Jossey-Bass.

Molnar, B.E., Miller, M.J., Azrael, D., & Buka, S.L. (2004). Neighborhood predictors of concealed firearm carrying among children and adolescents: Results from the Project on Human Development in Chicago Neighborhoods. *Archives of Pediatrics & Adolescent Medicine, 158*(7), 657–664.

Morenoff, J.D., & Sampson, R.J. (1997). Violent crime and the spatial dynamics of neighborhood transition: Chicago, 1970–1990. *Social Forces, 76*(1), 31–64.

Mullan, F., Phillips, R.L., & Kinman, E.L. (2004). Geographic retrofitting: A method of community definition in community-oriented primary care practices. *Family Medicine, 36*(6), 440–446.

O'Campo, P. (2003). Invited commentary: Advancing theory and methods for multilevel models of residential neighborhoods and health. *American Journal of Epidemiology, 157*(1), 9–13.

Osypuk, T., & Sandro, G. (2007). Choosing appropriate levels to assess the relation between space and population health. In S. Galea (Ed.), *Macrosocial determinants of population health* (pp. 399–436). New York: Springer Media.

Raudenbush, S. (2003). The quantitative assessment of neighborhood social environments. In I. Kawachi & L. Berkman (Eds.), *Neighborhoods and health* (pp. 113–114). New York: Oxford University Press.

Sampson, R. (1997). Collective regulation of adolescent misbehavior: Validation results from eighty Chicago neighborhoods. *Journal of Adolescent Research, 12*(2), 227.

Sampson, R. (1999). What 'community' supplies. In R. Ferguson & W. Dickens (Eds.), *Urban problems and community development* (pp. 241–292). Washington, DC: Brookings Institution Press.

Sampson, R. (2001). How do communities undergird or undermine human development? Relevant contexts and social mechanisms. In A. Booth & N. Crouter (Eds.), *Does it take a village? Community effects on children, adolescents, and families* (pp. 3–30). Mahwah, NJ: Lawrence Erlbaum Associates.

Sampson, R. (2003b). Neighborhood-level context and health: Lessons from sociology. In I. Kawachi (Ed.), *Neighborhoods and health* (pp. 132–146). New York: Oxford University Press.

Sampson, R.J., Morenoff, J.D., & Gannon-Rowley, T. (2002). Assessing 'neighborhood effects': Social processes and new directions in research. *Annual Review of Sociology, 28*, 443–478.

Sampson, R.J., Raudenbush, S.W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science, 277*(5328), 918–924.

Sampson, R., & Wilson, W. (1995). Toward a Theory of Race, Crime, and Urban Inequality. In J. Hagan and R. Peterson (Eds.), *Crime and Inequality* (pp. 37–54). Stanford, CA: Stanford University Press.

Sastry, N., Pebley, A., & Zonta, M. (2002). *Neighborhood definitions and the spatial dimension of daily life in Los Angeles.* Los Angeles: RAND Corporation.

Shaw, C.R., & McKay, H.D. (1942). Juvenile delinquency in urban areas. Chicago: University of Chicago Press.

Tatalovich, Z., Wilson, J., Milam, J., Jerrett, M., & McConnell, R. (2006). Competing definitions of contextual environments. *International Journal of Health Geographics, 5*, 55.

Weiss, L., Ompad, D., Galea, S., & Vlahov, D. (2007). Defining neighborhood boundaries for urban health research. *American Journal of Preventive Medicine, 32*(6), S154-S159.

Xue, Y., Leventhal, T., Brooks-Gunn, J., & Earls, F. (2005). Neighborhood residence and mental health problems of 5–11-year-olds. *Archives of General Psychiatry, 62*(5), 554–563.