Geographic Information Systems

Co-Chairs: Erich M. Daub and Keith D. Harries

In 1853 a massive epidemic of cholera in London killed and disabled thousands of people. Therapy was unavailable; indeed, neither the mode of transmission nor the responsible agent had yet been identified. Despite these limitations, John Snow was able to contain the outbreak. Analyzing data which he carefully collected, Snow mapped the location of every known case; the clustering of cases around one public water pump was obvious. The rest of the story is public health history; the pump handle was removed and the epidemic rapidly subsided.

The story of the Broad Street pump has become a metaphor for public health—sound surveillance and careful epidemiology can often lead to successful prevention even absent a complete understanding of causal relationships. Snow appreciated the value of each of the three now-classic features of epidemiologic investigation: person, place, and time. More than anything else, Snow showed the power of place. Location spoke volumes. Absolute location conferred intrinsic environmental constraints on health; relative location revealed the "activity space" of daily life, with all its unique spatially-dependent risks. Today, after being largely neglected for seven generations, public health is slowly rediscovering the power of place.

Place matters. Geographers have recognized this since the inception of their discipline over two-thousand years ago. Some places are distinguished by their topography, others by their natural resources. Laws and regulations set one place apart from another. Culture, crime, climate, capital, civil unrest—every place has unique features. That these location-specific attributes may influence the incidence of disease and injury—as well as opportunities for intervention—is now awakening the public health community.

If geography has long recognized the importance of spatial variation on the human condition, why has public health taken so long to do the same? Much of the explanation for the current climate favoring a re-awakening to this approach—the "geographic approach"—is the recent proliferation of computer systems and software dedicated to manipulating and mapping spatial data. These systems, called *Geographic Information Systems (GIS),* are revolutionizing geography. They have the same potential within public health. The advent of these systems effectively puts into the hands of epidemiologists and prevention specialists the power to understand and manipulate "space" and "place". One consequence of this is that public health professionals must now learn to think geographically. As with many school children in the U.S. today, this knowledge is not yet common or easily acquired.

Geography has traditionally had four concerns: the characterization of places, the understanding of man-land relationships, accounting for spatial distributions, and the differentiation of areas and the formation of regions. Geographic analyses usually have one of two aims:

- To account for spatial variation
- To integrate all of the variation at a location so as to explain or characterize a place or region

Public health and injury control can utilize both of these aims to better understand location-specific influences of the environment (natural and built) and behavior, and to determine modifiable features of geographically variable phenomenon. GIS provides an important tool—a window into the geographic world.

GIS software ranges in its capabilities and cost. Some "desk-top" mapping programs are available at no cost, while other full-featured GIS packages cost several thousands of dollars (U.S.) and have long learning curves. Functional mapping and basic spatial analysis programs can be had for between $200 to $2000 (U.S.), and provide accessible methods by which to construct computer-generated maps and moderately sophisticated spatial analyses.

There are three required elements for the application of GIS to public health and injury control. First, a specific GIS software program must be acquired. These are readily available in most Asian and European countries and Australia;

less developed nations have benefited from United Nations-funded GIS software development. Secondly, computer-readable "boundary" files must be acquired. These are digital equivalents of hard copy maps showing various jurisdictional boundaries and perhaps related features (such as streets, topography, etc.). Boundary files will be unique to the GIS user's study area. In the U.S., one might acquire files for census tracts, postal zip-codes, and county or state boundaries, to name but a few. Other nations have their own geographies: postal zones (England), territories (Australia), enumeration districts (Sweden) are examples. These boundary files must be digitally compatible with the specific GIS software product acquired. While there are attempts to standardize spatial data and file structures within the U.S. and internationally, a plethora of frequently incompatible file types presently exists. The most expedient means of acquiring compatible boundary files is to purchase them directly from the manufacturer of the GIS software. Most GIS vendors produce a host of compatible boundary files for use with their product. Boundary files exist for most European and Asian nations; some African and South American countries also have well developed digital maps available. In the U.S. there are wealth of boundaries available for purchase from government and private sources. The third element required to apply GIS to public health and injury control is geographically-referenced attribute data. Attribute data is the raw data describing features of the population and/or "cases" to be mapped. Examples include death certificate files, trauma registries, census data, land and property data, and ambulance/EMS provider data, to name a very few. Such data is said to be "geo-referenced" if it includes at least one data field specifying some location-specific value for each record in the database. This might be a postal code, a county identifier, or (the penultimate geo-reference) a street address. GIS software assigns an X-Y ("latitude" - "longitude") coordinate to every record in the database based on each record's geo-referenced variable. Any record can have more than one geo-referenced variable, making analyses at finer geographic scales possible. The remaining attribute data serves as the basis for relational spatial-analytical operations performed by the GIS. The combination of multiple geo-referenced data sets is the hallmark of GIS. By "layering" data sets over one another, complicated spatial arrangements are easily identified and powerful spatial analyses are possible.

To illustrate these features, consider two examples drawn from work in the Baltimore, Maryland (U.S.A.) metropolitan area. The first concerns an analysis of 2,639 juvenile gun crimes during an 11-year period (190=80-1990) in Baltimore City. U.S. census data was used to develop a social stress index at the census tract-level. The 2,639 juvenile gun crime cases were "geocoded" (assigned x-y coordinates by a GIS system), and evaluated with respect to geographic patterns of social stress and selected demographic attributes. The GIS software was used to layer each of these three elements over one another, and to explore spatial relationships. Spatial patterns of juvenile gun crimes were noted in "high" stress neighborhoods. Other patterns revealed location-specific relationships between stress-intense areas and "victim-perpetrator" ethnicity; most "Black on Black" homicides were confined to "high" stress neighborhoods. Pronounced "edge" or "frontier" effects were seen at margins of differential stress neighborhoods.

A second example of the application of GIS to injury epidemiology and control is the Baltimore County Injury Prevention Program's integration of GIS into their established injury surveillance system. Hard-copy and electronic injury mortality data is obtained from the State Office of the Chief Medical Examiner. This data includes street address-level data relative to three locations for each case: the site of injury, the location of residence, and the location of death. GIS software layers this data with a variety of additional data sets: death certificate data, census data, hospital discharge data, EMS ambulance run reports, zoning data, liquor license data, and other data sets. Maps are constructed and spatial analyses performed to explore absolute and relative relationships of personal, behavioral, environmental, and institutional attributes to injury morbidity and mortality. For example, mapping and locational analyses of three-years of county resident homicide deaths revealed that 25 percent of county residents were injured (mostly with a firearm) outside the county boundary (e.g., within the city of Baltimore). Sixty percent were injured outside their home. The reverse was noted for suicide deaths: 70 percent were injured in their own home; 90 percent of these died there as well. GIS is introducing the "geographic approach" to injury control and public health. This approach re-establishes the importance of place. Geography lends itself especially to answering five types of questions:

1. Why is "X" distributed in a certain way, as opposed to all the other possible ways it could be distributed?

2. Why are there different rates at which "X" spreads over time through an area?

3. Why are there differences in the locational choices we make for our institutions or our interventions?

4. Why are there differences in direction or distance of movements of people, ideas, or phenomena?

5. Why are there differences in the images people hold about their communities and surrounding environment and how does this influence health or the prevention of disease and injury?

Participants in this Workshop deliberated on the accessibility, use, and expansion of GIS and geographic techniques to the study and control of injury. Several recommendations resulted:

- Compile communications illustrating the value and utility of GIS to public health and injury control.
- Make GIS and map analyses accessible to public health workers through training and educational forums.
- Inventory (nationally and internationally) the degree to which existing injury-related data sets contain (and could contain) geo-referenced data.
- Derive consensus and proscribe the acceptable level of spatial data to be captured by injury-related data systems.
- Study and promote ways to safeguard the confidentiality of geo-referenced public health data.
- Promote GIS mapping and analysis projects and sources of experience and guidance.
- Disseminate descriptions of available GIS software, boundary files, and geo-referenced injury-related data sets to national and international audiences.
- Improve the quantity and quality of geographic data available in medical and public health records.
- Create vehicles (task force, E-mail bulletin boards, newsletters, interest groups) to promote and disseminate GIS applications to injury control.
- Work on methods to determine appropriate denominators to use with geo-referenced injury data.
- Expand the querying of "place" to existing data collection mechanisms (e.g., hospital emergency/casualty departments, trauma centers, medical examiners, EMS, police).
- Promote "desk-top" mapping as an introduction to GIS techniques for public health and injury control professionals.

The "invisible college" of public health professionals introducing GIS into their epidemiologic and preventive routine is expanding. Concerted efforts need to focus on infusing GIS into injury prevention. A collection of interested public health professionals is forming to advance these goals and to work toward these recommendations. With the support of colleagues from other disciplines, and through the future efforts of the I.C.E. on Injuries, we hope to see the seeds first sewn almost 150 years ago reap a great harvest.