# Census Data for Health Preparedness and Response



**Jane L. Garb**
*Baystate Medical Center, USA*

**Richard B. Wait**
*Baystate Medical Center, USA*

## INTRODUCTION

The epidemiologist works with researchers in various disciplines as well as public and private health practitioners who are responsible for maintaining and improving the health of the population. Health is defined as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity"(WHO, 1948). The role of the epidemiologist is to provide information on the extent, etiology and risk of disease or injury. Key to this role is knowledge of the location, size and demographic makeup of the population at risk for developing disease or injury. By integrating this information with the location of cases as well as risk factors, epidemiologists can make a vital contribution to disease and injury prevention, intervention and response. This applies both to endemic or "usual" levels of both chronic diseases like cancer or heart disease and infectious diseases like pneumonia or influenza, and injuries like gunshot wounds or motor vehicle accidents. It also applies to disease epidemics or outbreaks like SARS; attacks by biological or chemical weapons such as Anthrax, ricin or sarin; and inadvertent natural or technological disasters including earthquakes, transportation disasters or widespread power interruptions. This chapter explores the types of census data for disease surveillance, prevention and intervention.

## BACKGROUND

### Sources of Geographic and Population Data

The decennial census (*Census 2000*, 2001) and American Community Survey (*American Community Survey*, 2003; *Meeting 21st Century Demographic Data Needs Implementing the American Community Survey: May 2002*, 2002) are the major sources of population data in a GIS. The census is conducted every 10 years. It provides a complete count of the U.S. population based on the census "short form" given to every individual. In addition, it provides detailed socioeconomic and demographic information for individuals, families, and households and selected information about housing, based on the "long form", given to a 1/6 sample of the population.

The American Community Survey (ACS) is an ongoing survey by the U.S. Census Bureau on an approximate 2.5% sample of the population, with over-sampling for small governmental units such as American Indian Reservations. Data is collected monthly and reported annually using detailed demographic and economic questions from the Census long form, with additional questions such as those on grandparents as care-givers. Currently the American Community Survey is limited to an annual survey of 700,000 households in 31 sites across country. However, full implementation will begin in 2004. The first data products for the smallest areas and populations will be available in 2009. Once data products are produced for a population group or area, they will be updated each year (*American Community Survey Operations Plan*, 2003). ACS survey quality measures are high, and these help compensate for somewhat smaller sample size and larger confidence intervals of the ACS compared to the decennial census long form (Griffin & Obenski, 2002).

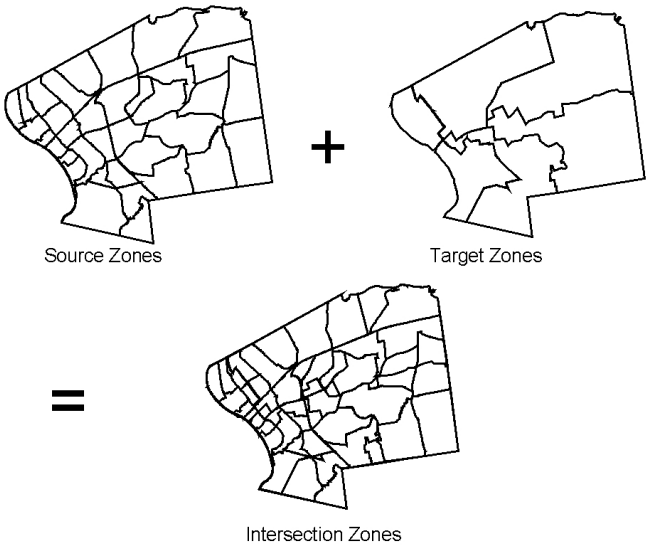
Census geographic data products are in the form of "TIGER"files (*TIGER/Line Files, Redistricting Census 2000*, 2001) which consist of street centerlines and address ranges as well as boundaries for census geographic units such as census tracts, block groups, cities and town boundaries, counties and regions.

All of these data products for the entire U.S. are publicly available from the Census Bureau Web site (http://www.census.gov) or on CD for a minimal price.

### Daytime Population Estimates

The ACS holds the most promise for current population estimates because of its ongoing nature. However, other sources of daytime population estimates show utility when combined with additional data. The Census Transportation Planning Package (CTPP) (*Census Transportation and Planning Package (CTPP)*, 2003) by the Department of Transportation is a special tabulation of Census

*Figure 1. Components of areal interpolation*




2000 data (long form responses) that includes information according to place of work, place of residence and travel between home and work. Discussions are underway about CTPP products that would be based on ACS data.

Another survey is the Longitudinal Employer - Household Dynamics survey (LEHD) (*Longitudinal Employer - Household Dynamics*, 2003), based on a partnership between the Census Bureau and currently 10 states. This survey integrates state administrative data and census data products. It includes place of residence and place of work for workers by age, sex, and industry. Also an ongoing survey, it has ties to the ACS, which will improve its accuracy and potentially provide extremely accurate estimates of daytime populations in the future. Neither of these latter two surveys is fully implemented at this point, and this will depend on future congressional funding.

## Data Integration in a GIS

We can integrate census data with health data in a GIS for prevention, intervention and response to diseases in all forms. This can be done for *endemic* or usual levels of chronic diseases like cancer or infectious diseases like pneumonia or influenza; infectious disease *epidemics* or outbreaks like the recent SARS epidemic; or sudden *attacks* with biological or chemical weapons. Census data allows us to perform critical GIS functions essential to mapping and analysis of spatial data in health.

To begin with, census geographic TIGER files are used as reference files for geocoding case or event locations. Census boundaries in these files are used in data aggregation of geographic areas and in the GIS function known as polygon overlay, where two map layers are overlaid on one another. This is the basis for a critical method necessary for utilizing census data for health called areal interpolation (Goodchild & Lam, 1980).

Both ACS and decennial census data are available only in aggregate form for fixed geographical units. These units – census tracts, block groups, counties – are administrative units used by the census to aid in census data collection, or by political entities to apportion representation and govern. However, the most useful geographic units for health applications, in which data are often available – neighborhoods, school districts, hospital service areas, and so forth – do not coincide with the census geography for which population counts are available.

Since 9/11, the scope of health has extended from preparedness against naturally occurring diseases and chronic conditions to biochemical weapons. The unique geography required for using census data in routine healthcare is even more of a challenge in the case of a bioterrorist attack, where the appropriate geographical units – police sectors, drainage sub-basins or the plume resulting from a chemical or biological release – may be unpredictable, and different depending on the nature of the biochemical weapon used.

Areal interpolation (AI) becomes a key method for addressing the challenge of estimating the population for the geography of healthcare. It starts with original geographic units for which data such as population counts is