Measuring Neighborhood Effects and the Use of Geo-coded Variables

Ninez A. Ponce, MPP, PhD
Associate Professor, UCLA Fielding School of Public Health
Associate Director, Asian American Studies Center
PI, California Health Interview Survey

CTSI Clinical Research Development Seminar
January 2013
Agenda

- Measuring Neighborhood Effects
- Geocoding
- Resources
Methodological challenges

- What’s the most appropriate level of geography?
- Can we accurately define neighborhood boundaries?
- Which characteristics of the social and physical environment are most relevant for health?
- How do we measure neighborhood characteristics?
- How do we parse out the relative influence of neighborhood and individual characteristics?
Unit?

- Obstacles to the use of area-based socioeconomic measures are both technical and conceptual.
- No consensus in the United States regarding which area-based measures should be used, at which level of geography, to measure or monitor socioeconomic inequalities in health.

Census block group (average population = 1,000; 600-3,000 people)
Census tract ("optimal" population = 4,000; 1,200-8,000 people)
US Postal Service zip code (large variation: about 1K to over 100K)

Krieger N et al. 2002
Geocoding System

Geocode Search Result for 2012 HMDA/CRA Reporting

<table>
<thead>
<tr>
<th>Street Address</th>
<th>1239 FERNWOOD PACIFIC DR</th>
<th>MSA/MD Code</th>
<th>31084</th>
</tr>
</thead>
<tbody>
<tr>
<td>City Name</td>
<td>TOPANGA</td>
<td>State Code</td>
<td>06</td>
</tr>
<tr>
<td>State Abbreviation</td>
<td>CA</td>
<td>County Code</td>
<td>037</td>
</tr>
<tr>
<td>Zip Code</td>
<td>90290</td>
<td>Tract Code</td>
<td>8001.02</td>
</tr>
</tbody>
</table>

MSA/MD Name: LOS ANGELES-LONG BEACH-GLENDALE, CA
State Name: CALIFORNIA
County Name: LOS ANGELES COUNTY
Empirical evidence that both choice of measure and level of geography matter

- Census block group and census tract measures performed similarly for virtually all outcomes.
- Zip code measures, however, in some cases failed to detect gradients or detected gradients contrary to those observed with the block group and tract measures.
- Categories based on quintiles and a priori cutpoints detected similar socioeconomic gradients, but only the latter could be uniformly applied across levels of geography within and across states.
- Economic deprivation (% poverty, Townsend index) measures were more robust than measures of education and wealth not only for leading causes of death and cancer, but also for deaths due to HIV and homicide.
Agenda

- Neighborhood effects
- Geocoding
- Resources
Geocoding

Arline T. Geronimus and John Bound, AJE 1999

- First employed in US health studies in the 1930s, the use of such geosocial measures—empirically observable social and physical characteristics of areas whose spatial distribution is patterned by human activity—facilitated by geographic information systems (GIS)

- Basic approach is to classify people in public health databases and in the total population by the socioeconomic characteristics of their residential neighborhood, using US Census

- These area-based geosocial measures—conceptualized as meaningful indicators of socioeconomic context in their own right and not merely "proxies" for individual-level data—can be validly applied to all persons, regardless of age, gender, and employment status
What is....

**Geocode**

*Geospatial Entity Object Code*

a representation format of a geospatial coordinate measurement used to provide a standard representation of an exact geospatial point location at, below, or above the surface of the earth at a specified moment of time (Wikipedia)

Can include some or all of the following geospatial attributes: Geocode Format Registry Number; Latitude; Longitude; Altitude; Others

**Geocoding**

the assignment of a code -- usually numeric -- to a geographic location, i.e., affixing to an individual address its latitude and longitude (Harvard, The Public Health Disparities Geocoding Project)

*Healthy People 2010 sets the goal of geocoding, by the year 2010, 90 percent of "all major national, state, and local health data systems… to promote nationwide use of geographic information systems (GIS) at all levels"*
Geographic Hierarchy for the 2010 Decennial Census (1)

* Refer to the "Hierarchy of American Indian, Alaska Native, and Native Hawaiian Areas" on page 2.
Geographic Hierarchy for the 2010 Decennial Census (2)
Online Cartographic & Geographic Resources

http://www.census.gov/geo/

Geography

About Us
- Who We Are
- What We Do
- Contact Us

Maps & Data
- Shapefiles
- KMLs
- Maps
- Gazetteers

Reference
- Definitions
- Concepts
- Codes

Partnerships
- Partner Programs
- Materials
- Schedules

Education

Research

Spotlight Product
2010 Population Distribution in the United States and Puerto Rico
More Info | Purchase

Trivia...
True or False: American expatriates are counted during the decennial census.
See the answer
## Examples of Available Census Data

**Social:**
- Household Type
- Marital Status
- Fertility
- Educational Attainment
- Veteran Status
- Disability Status
- Place of Birth
- Citizenship Status
- Language Spoken at Home
- Ancestry
- Linguistic Isolation

**Economic:**
- Employment Status
- Commuting to Work
- Occupation
- Industry
- Income
- Percent of Families/People below poverty level

**Housing:**
- Occupancy
- Housing Characteristics
- Housing Tenure
- Vehicles
- Heating Fuel
- House Value
- Mortgage Status
- Rent

**Demographic:**
- Total Population
- Gender
- Race
- Age

Source: [www.census.gov/2010census](http://www.census.gov/2010census)
Examples of Available Census Data

Measures of Inequality, Segregation, Exposure (a few examples)

Gini Coefficient
• measures dispersion of shares of aggregate income received by households, \( \textit{ranges from} 0 \) (complete equality) to 1 (complete inequality)

Dissimilarity index
• measures the percentage of a group's population that would have to change residence for each neighborhood to have the same percent of that group as the larger area overall., \( \textit{ranges from} 0 \) (complete integration) to 1 (complete segregation)

Information index or Entropy Index
• measures the (weighted) average deviation of each areal unit from the metropolitan area's racial and ethnic diversity)

Isolation index
• measures the extent to which minority members are exposed only to one another

Source: http://www.census.gov/hhes/www/housing/housing_patterns/app_b.html
Examples of Available Data from ESRI

Demographic
Population, households, housing, occupancy, income, age, race, Hispanic origin, and Census 2010 Data

Crime Risk
Major personal and property crime categories such as murder, rape, robbery, assault, burglary, theft, and motor vehicle theft

Community Information
demographic data, business information, and spending data for various sectors including: Banking and financial services, Education, Health and Human Services, Other

Consumer Data
Total Expenditures, Average Spending Per Household, and a Spending Potential Index (SPI)

Business Data
Total number of businesses by industry classification, Total sales, Total number of employees

Source: www.esri.com
Linking Community Level Data to Individuals in a Data Set

Example CHIS & using STATA

CHIS Source data (i.e. not the public use data) has address, census block, census tract, zip code, count information, but you must access through the Data Access Center (DAC) and requires a formal request to obtain permission.

Merge data by unit, i.e. tract

```stata
.use chis2009.dta
.sort tract
.save, replace
.use mycensusdata.dta
.sort tract
.merge tract using chis2009
.tab merge
.save chis2009_census
```
Analyzing changes in health inequalities through space and time

- Change in outcomes

- Components of change:
  - Compositional – change in population – race, ethnicity, age, income
  - Contextual – social, environmental, policy changes

- Statistical methods (next slide and next lecture)
Measuring Effects of Place on Health

- Multilevel methods can look at the health of neighborhoods after controlling for the health and other characteristics of individuals
  - compositional factors
    - the characteristics of people in particular places,
  - contextual factors
    - opportunity structures in the local environment such as access to food and transportation resources, and
  - collective factors
    - sociocultural and historical features of neighborhoods

- Next time: crash course on multilevel modeling
Agenda

- Neighborhood effects
- Geocoding
- Resources
Neighborhood data and resources for instruments

- Project on Human Development in Chicago Neighborhoods
  www.icpsr.umich.edu/PHDCN/instruments.html

- Community Tracking Survey
  www.hschange.com/index.cgi?data=01

- Area Resource File
  http://arf.hrsa.gov/

- County Business Patterns Database (voluntary associations/zip code)
  http://www.census.gov/econ/cbp/

- USA Counties Database (voting patterns)
  http://censtats.census.gov/usa/usa.shtml

- Neighborhood Change Database (tract)
Geocoded datasets for research on health

California Health Interview Study (CHIS)
Los Angeles County Health Survey (LACHS)
The Los Angeles Family and Neighborhood Survey (L.A.FANS)
MultiEthnic Study of Atherosclerosis (MESA)
Atherosclerotic Risk in Communities (ARIC)
Cardiovascular Health Study (CHS)
Hispanic Community Health Study- Study of Latinos (HCHS- SOL)
Translating Research into Action for Diabetes (TRIAD)
National Health and Nutrition Examination Study (NHANES)
Jackson Heart Study (JHS)
Look AHEAD (Action for Health in Diabetes)
Web tools for mapping data

healthycity.org
A free online mapping interface that includes a wide variety of indicators from the U.S. Census, American Community Survey, and the California Health Interview Survey – California data
Example (using healthycity.org):
Voter participation for a particular census tract
References


Ecob R, Macintyre S. Small area variations in health-related behaviors; do these depend on the behavior itself, its measurement, or on personal characteristics? Health and Place. 2000;6:261-274.


