How geographic data can enhance your administrative data research

Mario Schootman, Ph.D.
Associate Professor
Washington University School of Medicine

November 16, 2011
Outline

- How does geography relate to biomedical research?
- What can a GIS do for me? Why do I care?
- Where can I find geographic data?
- How do I analyze these data?
- What resources are available to incorporate geography into my research?
How does geography relate to biomedical research?
How does geography relate to biomedical research?
Fig. 1. Proposed causal indirect-cognitive and direct-contextual paths linking place to health.
What can a GIS do for me?

- Health and disease
- Social Factors
- Biodiversity
- Engineering
- Land Use
- Environmental Considerations
Functions of a GIS

- Input
- Storage
- Manipulation
- Query
- Analyze
- Visualization
Query & Analysis

- Evaluate traffic flow
Query & Analysis

Query & Analysis

- Evaluate areas most susceptible to landslide
Query & Analysis

Planning for a clinic location
Query & Analysis

New Hampshire Primary Care Service Areas
Children Under 18 Per Clinically Active Pediatrician

Data Source: U.S. Census, 2000
AMA/ACr MapArt/Inc, 2000

Legend:
- fewer than 1000 (8)
- 1000-2000 (7)
- 2000-4000 (9)
- 4000 or more (4)
- no clinically active pediatrician (13)

Non-poza, rural lands are shown in gray.

David Goodman, MD FAAP
The Primary Care Service Area Project
The Center for Evaluative Clinical Sciences
Dartmouth Medical School
- Multidisciplinary
- Integrated
- Holistic
Geography Can Integrate Data

- Visualizing
- Connecting
- Relating
Where can I find geographic data?
Where can I find geographic data?

Geocoding of 107 Oak Avenue
Where can I find geographic data?

![Image of HRSA website](image)

**HRSA** U.S. Department of Health and Human Services
Health Resources and Services Administration

**Area Resource File (ARF)**
National County-level Health Resource Information Database

**Looking for **SPECIFIC** health-related data?**
Search the ARF

**2009-2010 ARF Release is Now Available.**

**Order the ARF on CD or download now >**

New webtool - HRCT
A Health Resources County Comparison Tool is now available comparing poor counties on population characteristics and health resources.

**Sources**
The ARF is a collection of data from more than 50 sources, including:
- American Medical Association
- American Hospital Association
- US Census Bureau
- Centers for Medicare & Medicaid Services
- Bureau of Labor Statistics
- National Center for Health Statistics

**Users**
The ARF is designed to be used by planners, policymakers, researchers, and others interested in the nation's health care delivery system and factors that may impact health.

---

**Health Care Professions**
- Physicians by detailed specialty and major professional activity, gender
- Dentists
- Physician assistants
- Nurse practitioners
- Data for other major health professions

**Health Facilities**
- Services
- Beds by type
- Hospital personnel
- HIV/AIDS
- Nursing facilities
- Ambulatory surgery centers
- Hospitals
- Home health agencies

**Health Professionals Training**
- Statistics on Medical, Pharmacy, Dental, Optometry and other health professions training schools
- Schools by County
- Enrollments
- Graduate levels

**Census, Population Data and Environment**

**Hospitals and Health Care Facilities**

**Population and Economic Data by County**
- Demographics by age, race, sex
- Employment
- Income levels
- Housing statistics
- Vital statistics
- Health insurance

**Environment**
- Population Per Square Mile
- Land Area

**Codes and Classifications**
- Metropolitan Statistical Areas
- EOS (Economic Census)
- Rural/urban codes
- Region codes
- Exchange area

---

18
Where can I find geographic data?

Research Data Assistance Center

Available CMS Data > Medicare > Data Available

Medicare Data Available

Identifiable | Limited | Non-Identifiable

*** CMS Data Availability Table ***

Research Identifiable Files (RIFs):

The RIFs contain person-specific data on Medicare providers, beneficiaries, and recipients including individual identifiers that would permit the identity of a beneficiary or physician to be deduced (e.g., date of birth, age, race, sex, residence information).

Data with beneficiary or physician identifiers are subject to the Privacy Act, Freedom of Information Act and other Federal government rules and regulations. As such, the information is confidential and is to be used only for reasons compatible with the purpose(s) for which the data are collected.

CMS employs strict security measures to safeguard individual privacy. CMS data release policies seek to ensure that files containing physician and/or beneficiary identifiers are used only when necessary and in accordance with disclosure provisions of the Privacy Act.

Researchers need to submit to CMS a data request packet to CMS's Privacy Board (after required ResDAC review). If CMS approves the release of the data, researchers need to pay the cost incurred in the processing of data.

RIF Data Available

Limited Data Set (LDS)

The LDS data contains beneficiary level health information but exclude specified direct identifiers as outlined in the Privacy Rule. The LDS data are considered identifiable data sets even without the specified direct identifiers. Because the information is considered identifiable, it remains subject to the Privacy Act.

To obtain the LDS data, data requestors must show CMS that their proposed use of the data meets the disclosure provisions for research purposes as defined in both the Privacy Rule and...
Databases
The HCUP family of administrative longitudinal databases contains encounter-level information on inpatient stays, emergency department visits, and ambulatory care in U.S. hospitals. These are available through AHRQ through a Federal-State-Industry partnership.

Nationwide HCUP Databases
Nationwide HCUP databases can be used to identify, track, and analyze national trends in health care utilization, access, charges, quality, and outcomes.

- Overview of the NIS
- How to Obtain the NIS
- HCUP Partners in the NIS
- Related Reports
- Database Documentation
- HCUP Data Use Agreement Training
- NIS Data Use Agreement (PDF file, 42 KB; or HTML)
- HCUP Supplemental Files
- Requirements for Publishing with HCUP Data

Kids' Inpatient Database (KID)

- Overview of the KID
- How to Obtain the KID
- HCUP Partners in the KID
- Related Reports
- Database Documentation
- HCUP Data Use Agreement Training
- KID Data Use Agreement (PDF, 42 KB; or HTML)
- HCUP Supplemental Files
- Requirements for Publishing with HCUP Data

Nationwide Emergency Department Sample (NEDS)

- Overview of the NEDS
- How to Obtain the NEDS
- HCUP Partners in the NEDS
- Related Reports
- Database Documentation
- HCUP Data Use Agreement Training
- NEDS Data Use Agreement (PDF, 41 KB; or HTML)
- HCUP Supplemental Files
- Requirements for Publishing with HCUP Data

State-Specific HCUP Databases
State-specific HCUP databases can be used to investigate State trends in health care utilization, access, charges, quality, and outcomes.

State Inpatient Databases (SID)

- Overview of the SID
- How to Obtain the SID
- HCUP Partners in the SID
- Related Reports
- Database Documentation
- HCUP Data Use Agreement Training
- SID Data Use Agreement (PDF, 41 KB; or HTML)
- HCUP Supplemental Files
- Requirements for Publishing with HCUP Data

State Ambulatory Surgery Databases (SASD)

- Overview of the SASD
- How to Obtain the SASD
- HCUP Partners in the SASD
- Related Reports
- Database Documentation
- HCUP Data Use Agreement Training
- SASD Data Use Agreement (PDF, 41 KB; or HTML)
- HCUP Supplemental Files
- Requirements for Publishing with HCUP Data

State Emergency Department Databases (SEDD)

- Overview of the SEDD
- How to Obtain the SEDD
- HCUP Partners in the SEDD
- Related Reports
- Database Documentation
- HCUP Data Use Agreement Training
- SEDD Data Use Agreement (PDF, 41 KB; or HTML)
- HCUP Supplemental Files
- Requirements for Publishing with HCUP Data
The Dartmouth Atlas of Health Care

Understanding of the Efficiency and Effectiveness of the Health Care System

For more than 20 years, the Dartmouth Atlas Project has documented glaring variations in how medical resources are distributed and used in the United States. The project uses Medicare data to provide information and analysis about national, regional, and local markets, as well as hospitals and their affiliated physicians. This research has helped policymakers, the media, health care analysts and others improve their understanding of our health care system and forms the foundation for many of the ongoing efforts to improve health and health systems across America. LEARN MORE

PERCENT OF PATIENTS READMITTED WITHIN 30 DAYS OF DISCHARGE FOLLOWING A MEDICAL ADMISSION

http://www.dartmouthatlas.org/
SEER-Medicare Linked Database

- **About the SEER-Medicare Database**
  - Brief Description of the Database
  - How the SEER & Medicare Data are Linked
  - Publications Using SEER-Medicare Data
  - Fact Sheet (PDF)

- **The SEER-Medicare Data Files**
  - About the Data Files
  - SEER Program & Data
  - Medicare Enrollment & Claims Data
  - Summary Table of Available Medicare Data
  - Provider Files
  - File Sizes & Distribution
  - Number of Cases for Selected Cancers
  - HCPCS Tables
  - Frequency of NDC Codes on Durable Medical Equipment Files
  - Number of Part D Enrollees
  - Frequency of NDC Codes on Part D Event Files
  - Potential Funding for SEER-Medicare Analyses

- **Obtaining the SEER-Medicare Data**
  - Overview of the Process for Obtaining the Data
  - Instructions & Data Use Agreements
  - Instructions for Requesting New Data for Previously Approved Projects
  - Proposal Review Process
  - Requirements Following Receipt of Data
  - Cost of Acquiring SEER-Medicare Data
Breast cancer cluster of higher risk of death (cluster 1),
Lower risk of death (cluster 2),
and outside cluster areas with census-tract poverty rate for women with breast cancer age 66 or older diagnosed in 1992-1999, Detroit, MI.
Breast cancer cluster of higher risk of death (cluster 1), lower risk of death (cluster 2), and outside cluster areas with census-tract African Americans for women with breast cancer age 66 or older diagnosed in 1992-1999, Detroit, MI.
How do I analyze these data?

- Multilevel model
- Spatial model
- Small areas & unstable rates
- Availability of geographic areas
What resources are available?

- GIS lab at WU: http://gis.wustl.edu/GIS/Resources.html
  - Training
  - Software
  - Great people!

- CADR (http://cadr.wustl.edu/)

- Health Behavior Communication & Outreach Core (Siteman Cancer Center & CADR)
What resources are available?

- Many, but their use depend on your needs
- Data ...
  - About neighborhoods & locations of medical care
    - Observed conditions
    - Perceived conditions
    - Existing data (e.g., census, Google Earth)
- Travel through space & time
geographic data can enhance your administrative data research