Geographically Discrete Data: Mining the Internet to Find Authoritative Health Data

Helena VonVille

Library Director

The University of Texas School of Public Health at Houston

Goal of CHARTing Web site

- Provide a comprehensive collection of links
  - Geographically discrete data for Texas
  - Illness/disease incidence/prevalence
  - Describe determinants of health
- Select from authoritative data providers
- Site stability important
- Actively update site

Determinants of Health

o Income and Social Status
o Social Support Networks
o Education
o Employment/Working Conditions
o Social Environments
o Physical Environments
o Personal Health Practices and Coping Skills
o Healthy Child Development
o Biology and Genetic Endowment
o Health Services
o Gender
o Culture

Geographically Discrete Data: Definition

- Also known as "localized data"
- Describes a locale at the county, city, or neighborhood level
- Dependent upon the needs of person using the data
  - Researchers in urban areas may need more discrete data than those in rural areas

Community Health Assessment Resource for Texas (CHART

- Funded by NN/LM SCR
- Geocode Texas vital statistics and Census data
- GIS-based site would be publicly available
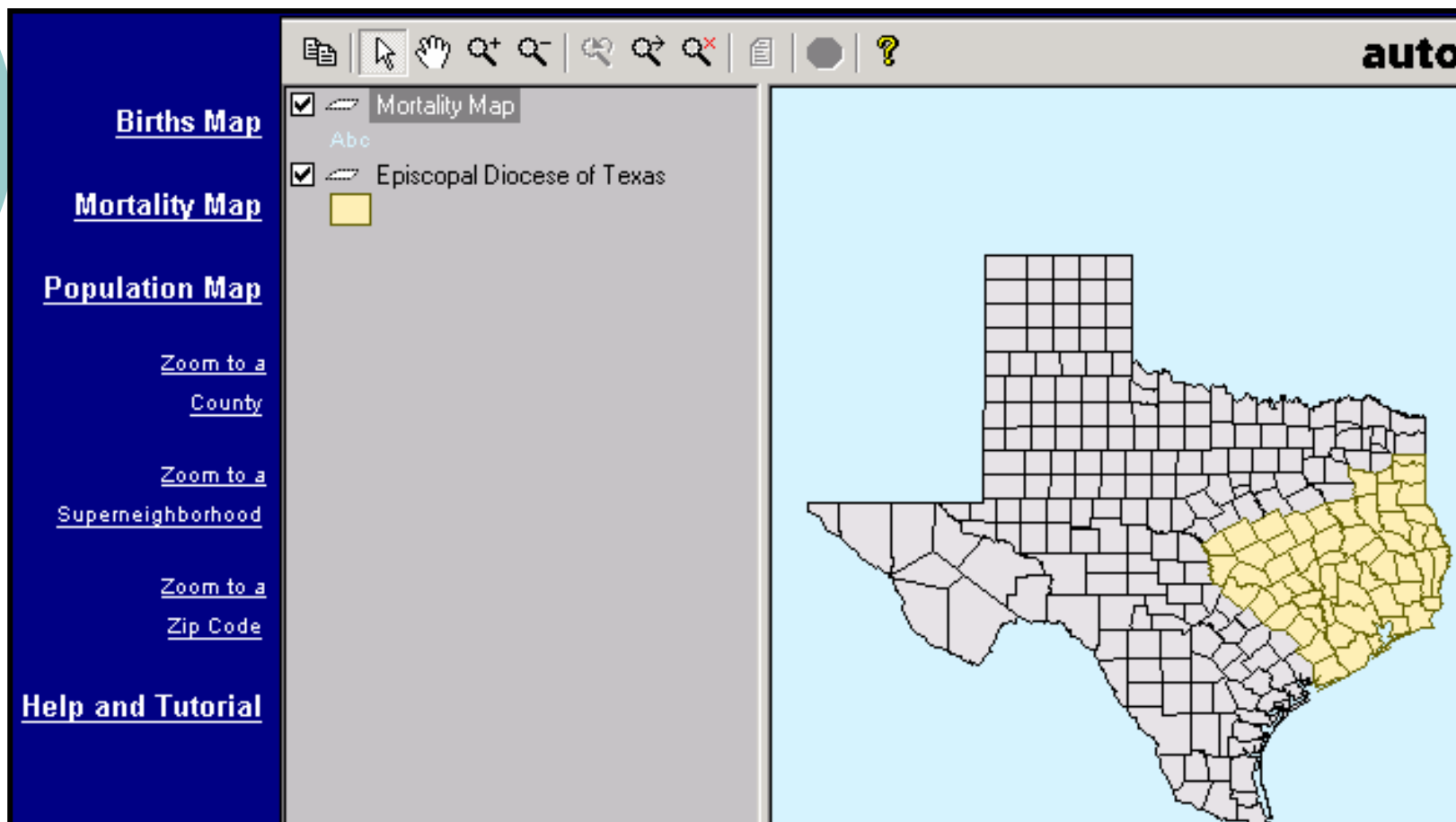- Library role: dissemination and education
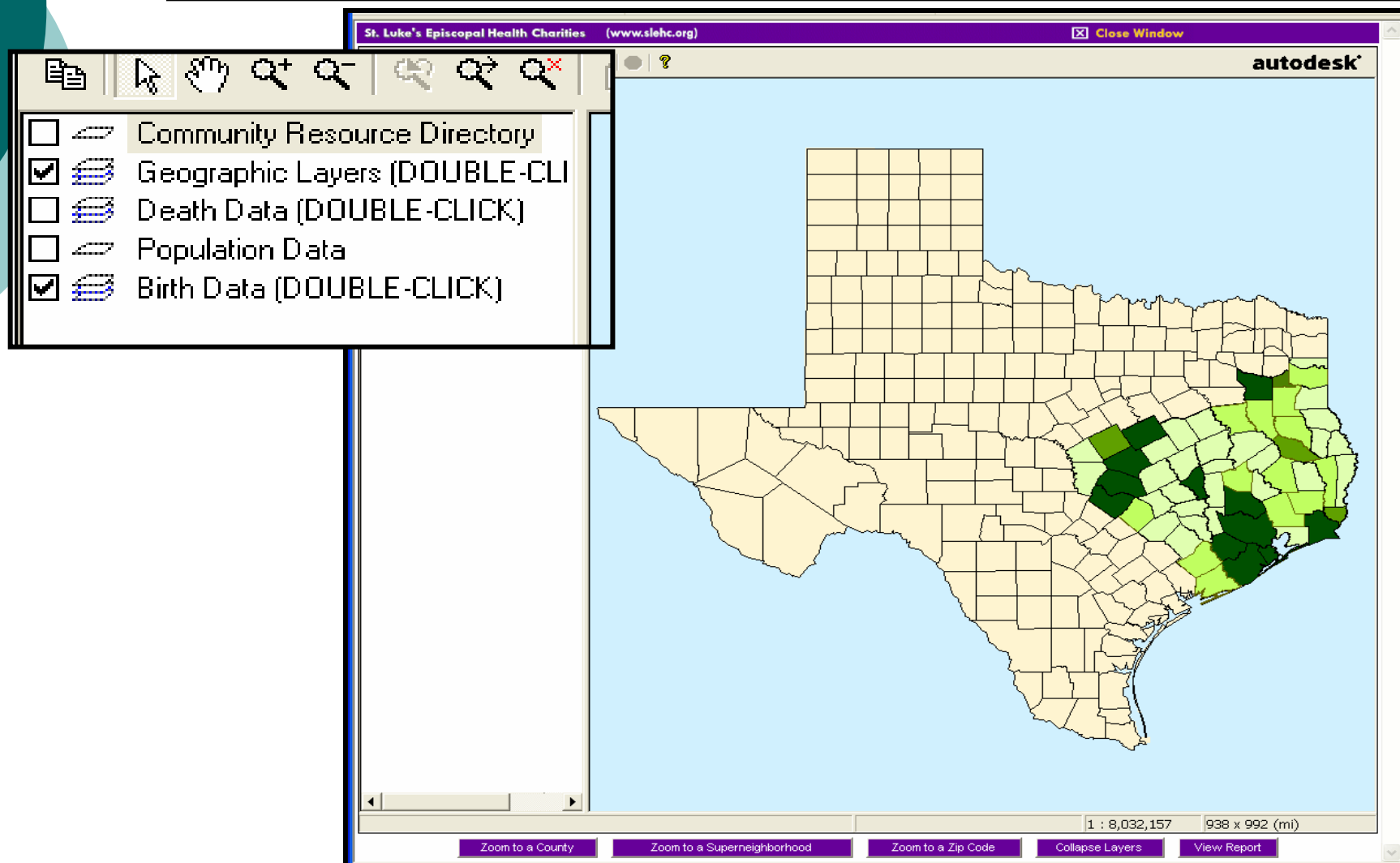
Problems with Geocoding Data

- Incorrect address on birth/death certificates
- Incorrect county on birth/death certificates
- Street names misspelled
  - MacGregor vs. McGregor
- Numbers transposed or read incorrectly during input process
  - 4 that looks like a 9
- Post office boxes used in rural areas
  - 80% to 90% of urban data generally matched on first sweep
  - 50 to 65% of rural data can be initially matched
- Texas is really big

Community Health Information System (CHIS) (2003)

Community Health Information System (CHIS) (2004)

Original CHARTing Menu

- Glossary of terms
- County overviews
- Natality, mortality, and morbidity data
- Sociodemographic data
- Environmental & occupational exposure data

County Overviews

Texas Health Facts (formerly County Fact Sheets; Texas Department of Health; 199
Good 2 page overview of the health of a **county** that includes: Demography, Socid
Indicators, Medicaid, Natality, Communicable Diseases, Mortality, Health Profession
Facilities, and Texas Department of Health Program Summaries. Scroll down for ear

The Landscape Project (Texas Institute for Health Policy Research; **county** level da
Metropolitan Statistical Area level data)
The Landscape Project creates reports based on 14 indicators: Communicable Dise
Environmental Health, Government Finance, the Health Care Sector, Health Insurar
Information, Infant and Maternal Health, Mortality Statistics, Needs Based Program
Projections, Populatin Distribution, Social Security, and Socioeconomic Characterist

Community Profiles (Texas Economic Development; **city** and **county** level data)
The Community Profiles give very brief overviews of each county and include demo
educational enrollment, farming statistics, public assistance expenditures and pove
data.

Changes to CHARTing Menu

- Replaced Morbidity, Mortality, Natality with:
  - Infectious Diseases
  - Maternal Health, Childbirth, & Children
  - Chronic Diseases
  - Cancer
  - Cardiovascular Disease
  - Suicide, Accidents, & Injuries
  - Other Causes

Texas Department of Health

- Birth Defects
- Behavioral Risk Factor Surveillance
- Breast & Cervical Cancer Control
- Cancer Registry
- Center for Health Statistics
- Epidemiology
- HIV/STD, AIDS
- Infectious Disease Epidemiology and Surveillance

Morbidity Data from TDH

Public Health Region 8 Statistical Summaries

- Reported AIDS & HIV Infection Rates - 1999

- Reported Selected Gastrointestinal Disease Rates - 1999

- Reported Hepatitis Rates - 1999

- Reported Other Selected Disease Rates - 1999

- Reported Sexually Transmitted Disease Rates - 1999

- Reported Vaccine-Preventable Disease Rates - 1999

Morbidity Data from TDH

REPORTED OTHER SELECTED DISEASE RATES
(CASES PER 100,000 POPULATION)
PUBLIC HEALTH REGION 8 - 1999

COUNTY	1999 POP.	ASEPTIC MENINGITIS		CHICKENPOX		ENCEPHA
		CASES	RATES	CASES	RATES	CASES
ATASCOSA	36,915	1	2.7	1	2.7	0
BANDERA	13,915	0	0.0	3	21.6	0
BEXAR	1,360,411	147	10.8	2	0.1	1
CALHOUN	20,099	0	0.0	0	0.0	0
COMAL	73,767	1	1.4	31	42.0	0
DEWITT	20,314	0	0.0	11	54.1	0
DIMMIT	11,251	0	0.0	41	364.4	0
EDWARDS	2,497	0	0.0	0	0.0	0

Morbidity Data from TDH

1999 Annual Epidemiology Report (Texas Department of Health)

1998 Annual Epidemiology Report (Texas Department of Health)

1997 Annual Epidemiology Report (Texas Department of Health)
Go to **Regional Statistical Summaries** for reports by county for chicken pox, hepatitis, STD's, measles, mumps, etc. Go to **Reported Cases of Selected Diseases** for reports by public health regions.

CDC Wonder: Natality

Natality Data Request

Data Use Restrictions

...set has counts of births occuring within the United States to U.S. residents and non-residents. Counts can...
...by state, county, child's gender and weight, maternal race, maternal age, maternal education, gestation p...
...care, and birth plurality. The data is derived from birth certificates. More information.

...ve trouble displaying the list of states below, click here for the screen without java applets.

1. Select location and time span:

Select state, and then select county:
(Mother's State and County of Residence)

The United States

Year(s) (1995 - 2001): 1995 Through: 2001

2. Select maternal and child characteristics:

Gender of Child:	Any		
Birth Weight:	500 gms or less	through:	Not stated
Plurality of Birth:	Any		
Maternal Age:	Less than 15 years	through:	50-54 years
Maternal Race:	Any Races		
Hispanic Origin of Mother:	Any Hispanic and Non-Hispanic		
Maternal Education:	0 - 8 years	through:	Not stated
Prenatal Care Began:	No Prenatal care	Through:	Unknown or not stated
Gestation Period:	Under 20 weeks	through:	Not stated

CHARTing Link to Federal Reports

○ <u>Tuberculosis Surveillance</u> (Centers for Disease Control & Prevention; data available for **MSAs** and **selected cities**)
Scroll down to Tables 38 through 43 for TB morbidity by MSA and selected cities.

Landscape Project (Texas Institute for Health Policy Research)

- Utilizes data from CDC, TDH, FBI, and other agencies
- Works in collaboration with an advisory committee from state agencies, health care professionals and community representatives
- Provides reports for counties or MSAs
- Reports include 120 indicators in 14 categories

- Communicable Diseases
- Crime
- Environmental Health
- Government Finance
- The Health Care Sector
- Health Insurance
- Household Information
- Infant and Maternal Health
- Mortality Statistics
- Needs Based Programs
- Population Projections
- Population Distribution
- Social Security
- Socioeconomic Characteristics

Missing Sources for County-Level Data

- Non-reportable infectious diseases
  - Can find emerging disease data (SARS, WNV)
    - Extremely small incidence
- Chronic conditions
  - Asthma, diabetes, obesity, and cardiovascular disease
- Mental illness
- Current birth defects data only recently made available

Texas Hospital Inpatient Discharges

- Multiple report variables may be selected to create tables
- Utilizes DRGs, not the more specific ICD-9 or ICD-10
- Public Health Region level of data

Texas Hospital Inpatient Discharges

Hospital Discharge Statistics for Public Health Region 6

Year: 2000

Age of Patient: 0 to 17

Diagnosis: (4) Diseases and disorders of the respiratory system

Status	Payer																
	Self-Pay		Worker's Compensation		Medicare		Medicaid		CHAMPUS and other Federal Programs		Commercial/Blue Cross		Other		Unknown		Total	
	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay	Number	Total Length of Stay
Routine Discharge	1,011	3,079	3	8.0	14	34	4,346	18,860	33	81	3,704	13,586	166	914	55	133	9,332	36,695
Discharge to Another Facility	20	48	0	0.0	0	0.0	95	1,048	1	1.0	67	646	6	86	2	5.0	191	1,834
Discharge to Home Health Service	0	0.0	0	0.0	0	0.0	37	346	0	0.0	16	164	0	0.0	0	0.0	53	510
Left Against Medical Advice	3	9.0	0	0.0	2	2.0	16	29	0	0.0	4	11.0	0	0.0	1	3.0	26	54
Expired	0	0.0	0	0.0	0	0.0	25	423	0	0.0	14	207	5	19.0	0	0.0	44	649
Other	0	0.0	0	0.0	0	0.0	1	3.0	0	0.0	4	9.0	0	0.0	1	3.0	6	15.0
Unknown	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Total	1,034	3,136	3	8.0	16	36	4,520	20,709	34	82	3,809	14,623	177	1,019	59	144	9,652	39,757

Texas Hospital Discharge Database

2. Enter ICD-9 Code

`250` **ICD-9 Code LookUp**

Example: Enter "250.7" for All ICD-9 Codes that begin with "250.7".

Example: Enter "001.0" for "CHOLERA DUE TO VIBRIO CHOLERAE", the specific ICD-9 Code Description for "001.0".

3. Select Patient Gender
- Male
- Female
- Both Genders
- Unknown
- ⦿ All

4. Select Patient Age
- 35 - 39 years
- 40 - 44 years
- 45 - 49 years
- 50 - 54 years
- 55 - 59 years
- 60 - 64 years
- 65 - 69 years
- 70 - 74 years
- 75 - 79 years
- 80 - 84 years

5. Select Patient Race
- American Indian/Eskimo/Aleut
- Asian or Pacific Islander
- Black
- White
- Other
- ⦿ Any

6. Select Patient Eth
- Hispanic Origin
- Not of Hispanic O
- ⦿ Any

7. Select State/County/Zip Code of Patient Residence
Instructions on Selecting All Counties/Zip Codes

AR	Fort Bend	77053
LA	Franklin	77053
NM	Freestone	77406
OK	Frio	77417
TX	Gaines	77441
	Galveston	77444
	Garza	77451
	Gillespie	77459
	Glasscock	77461
	Goliad	77464

8. Select County/Hospital of Discharge in Te
Instructions on Selecting All Counties/Hospitals

TX	Hansford	BayCoast Medical Cen
	Hardeman	Bayou City Medical Ce
	Hardin	Bayou City Medical Ce
	Harris	Bayshore Medical Cen
	Harrison	Beacon Health Westch
	Haskell	Bellaire Medical Center
	Hays	Bellaire Medical Center
	Hemphill	Ben Taub General Hos
	Henderson	Casa A Special Hospit
	Hidalgo	Charter Behavioral Hea

○ Report variables include:
- Quarter/Year of discharge
- ICD-9 code
- Gender of patient
- Race/ethnicity
- Patient's state, county, zip code
- Hospital state, county, zip

Texas Hospital Discharge Database

- If 3-digit ICD-9 used, long list of 4- and 5-digit ICD-9's returned

- Final result at bottom of page

Developing Context: What's Wrong with This Data?

As you view these links to various sources of health data, it is important to keep in mind that data does not tell the whole story. Data and data gathering are not perfect-- if you see an anomaly or large deviation in the data, find out why. Don't assume.

1. Do you know what data is really gathered?

Births for Harris County								
	Marital Status							
	Married		Not Married		Unknown		All Mothers	
Year	Number	Percent	Number	Percent	Number	Percent	Number	Percent
1992	45,851	78.2	12,801	21.8	0	0.0	58,652	100.0
1993	*45,873*	*78.8*	*12,332*	*21.2*	0	0.0	58,205	100.0
1994	*39,130*	*68.6*	*17,691*	*31.0*	205	0.4	57,026	100.0
1995	38,208	67.0	18,783	32.9	66	0.1	57,057	100.0

(*Texas Health Data Births to Texas Residents*)

2004 Changes to CHARTing

- Partnership Links
  - National Network of Libraries of Medicine South Central Region
  - Community Health Assessment Resource for Texas
  - Texas State Strategic Health Partnership Data context
  - Texas Public Health Training Center

- New pages added ("Other Sources of Data")
  - Access to Health Care
  - Criminal Justice Data
  - Regional Web Sites

CHART Today

CHART Today

2001 Birth Report

Births by Trimester of Prenatal Care and Ethnicity

2001 Number of Births by Trimester of Prenatal Care and Ethnicity

COUNTY	1st Trimester Prenatal Care	2nd Trimester Prenatal Care	3rd Trimester Prenatal Care	No Prenatal Care	TOTAL
Milam	304	39	8	8	359
TOTAL	304 84.7%	39 10.9%	8 2.2%	8 2.2%	359 100.0%

2001 Number of Births by Trimester of Prenatal Care and Ethnicity

White

COUNTY	1st Trimester Prenatal Care	2nd Trimester Prenatal Care	3rd Trimester Prenatal Care	No Prenatal Care	TOTAL
Milam	179	20	0	5	204
TOTAL	179	20	0	5	204

Possible New Layout

Conceptual Framework

- Hillemeier, M.M., Lynch, J., Harper, S., and Casper, M. 2003. "Measuring Contextual Characteristics for Community Health." *HSR: Health Services Research* 38 (6, Pt. II): 1645-1717)

To visit the CHARTing site or to contact me about the CHARTing challenge:

http://www.sph.uth.tmc.edu/library/chartinghealthinfo.htm

Thank you!