National Cancer Institute

Collecting, Analyzing, and Presenting Geographic Information in Health Survey Data

David Stinchcomb[1], Jeremy Lyman[2], William W. Davis[1], and Richard P. Moser[1]

[1] National Cancer Institute, [2] IMS, Inc.

October 9, 2007

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

Introduction

- Geographic information in survey data
  - Respondent's residence, workplace, space-time paths, etc.
  - Ability to stratify results by geographic area
    - Region, state, county, etc.
  - Enable spatial data linkages (via GIS, e.g.)
- Three examples – collecting, analyzing, presenting:
  - California Health Interview Survey (CHIS)
  - Behavioral Risk Factor Surveillance System (BRFSS)
  - Health Information National Trends Survey (HINTS)

Example 1 – CHIS

- Collects respondent's residence location
- Hierarchical series of information sources:
  - Mailing address: "Is this where you live"?
  - Ask for street address of residence
  - Ask for nearest intersection
  - Use mailing address ZIP code

CHIS 2005 – Address Questions

QA05_N2 Your phone number was randomly selected for this study by a computer. We were able to match an address to your phone number to send a letter to your home explaining the purpose of this study.

Originally: "Is this your current address"

Do you now live at {R's address and street}?

YES.....1 **[GO TO QA05_N6]**
NO.....2
REFUSED.....-7
DON'T KNOW.....-8

QA05_N3 What is your zip code?

_____(ZIP CODE)
REFUSED.....-7
DON'T KNOW.....-8

QA05_N4 To help us better understand the environment you live in and how it may affect your health, please tell me the address where you live. This information will be kept confidential.

_____ (HOUSE ADDRESS NUMBER)
_____(NAME OF STREET, VERIFY SPELLING) **[GO TO QA05_N6]**
NO.....2
REFUSED.....-7
DON'T KNOW.....-8

QA05_N5 Can you tell me just the name of the street you live on?

_____(NAME OF STREET)
REFUSED.....-7 **[GO TO CLOSE1]**
DON'T KNOW.....-8 **[GO TO CLOSE1]**

QA05_N6 And what is the name of the street down the corner from you that crosses your street?

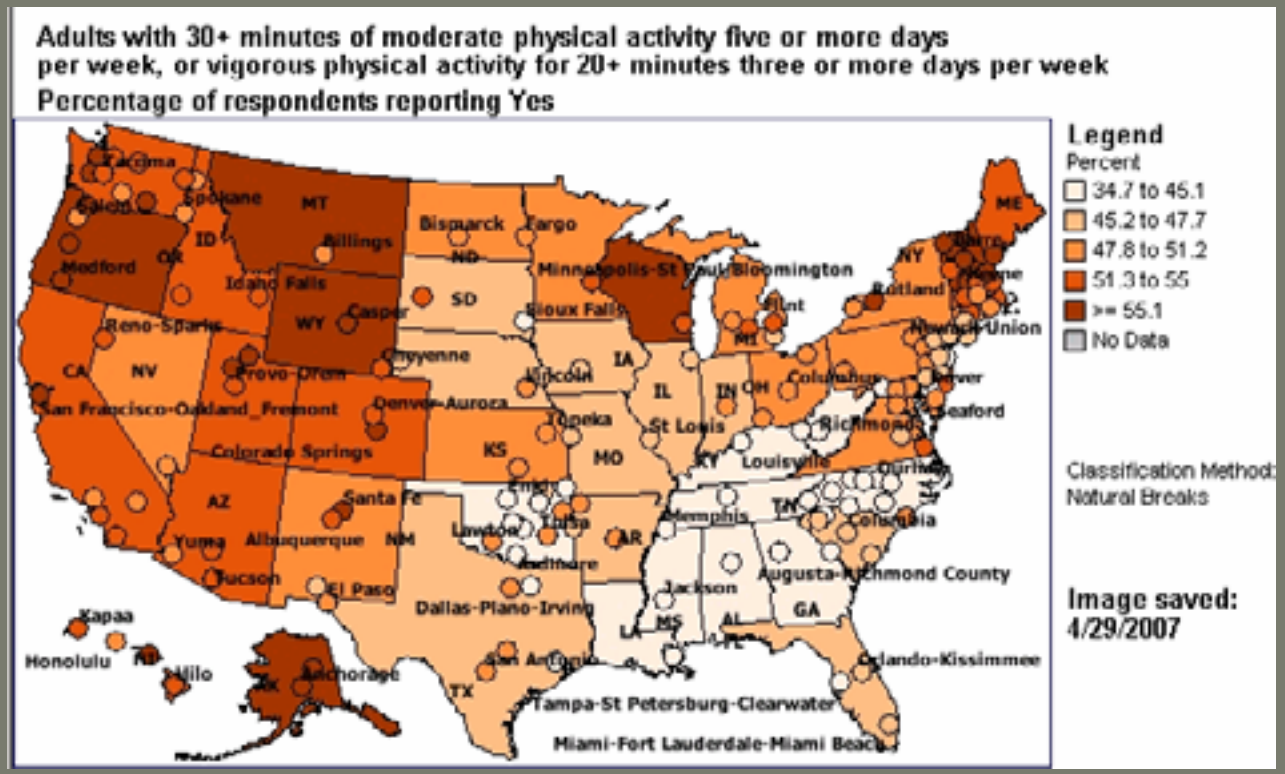
_____(NAME OF CROSS-STREET)
REFUSED.....-7
DON'T KNOW.....-8

CHIS Geocoding Rates

- Geocoding match rates (CHIS 2003):
  - Street address: 85.9%
  - 9-digit ZIP code: 1.1%
  - 7-digit ZIP code: 0.4%
  - 5-digit ZIP code: 12.6%

Example 2 – BRFSS

- Large national phone-based survey
- Results available by state and MSA:



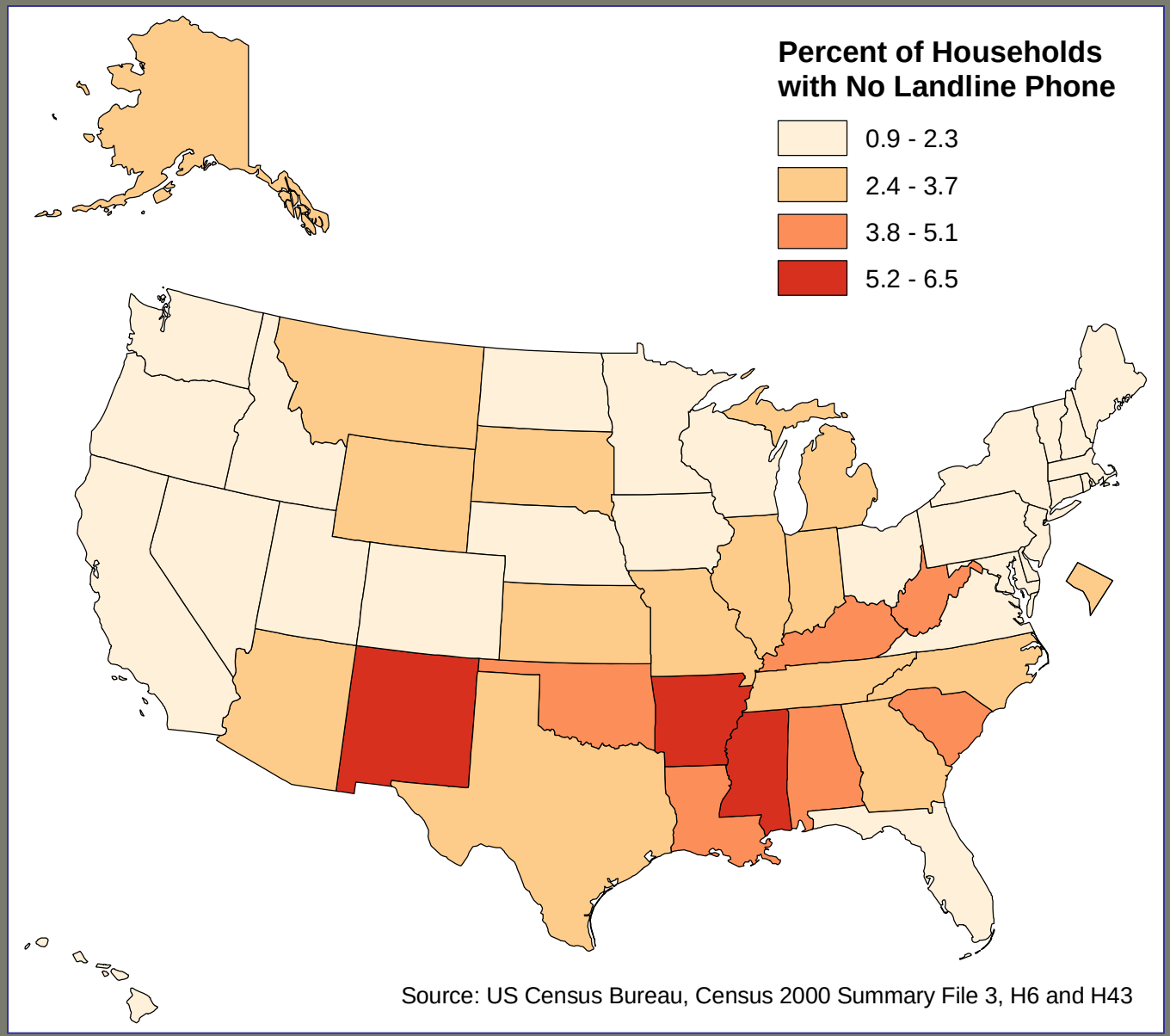
BRFSS with NHIS

- Statistical modeling project to combine BRFSS with the National Health Interview Survey (NHIS)
  - NHIS: large national in-person survey
- NCI, NCHS, Univ. of Michigan, Univ. of Pennsylvania
- Still in research phase
- Two main goals:
  - Correct for telephone-based survey bias
  - Improve geographic detail

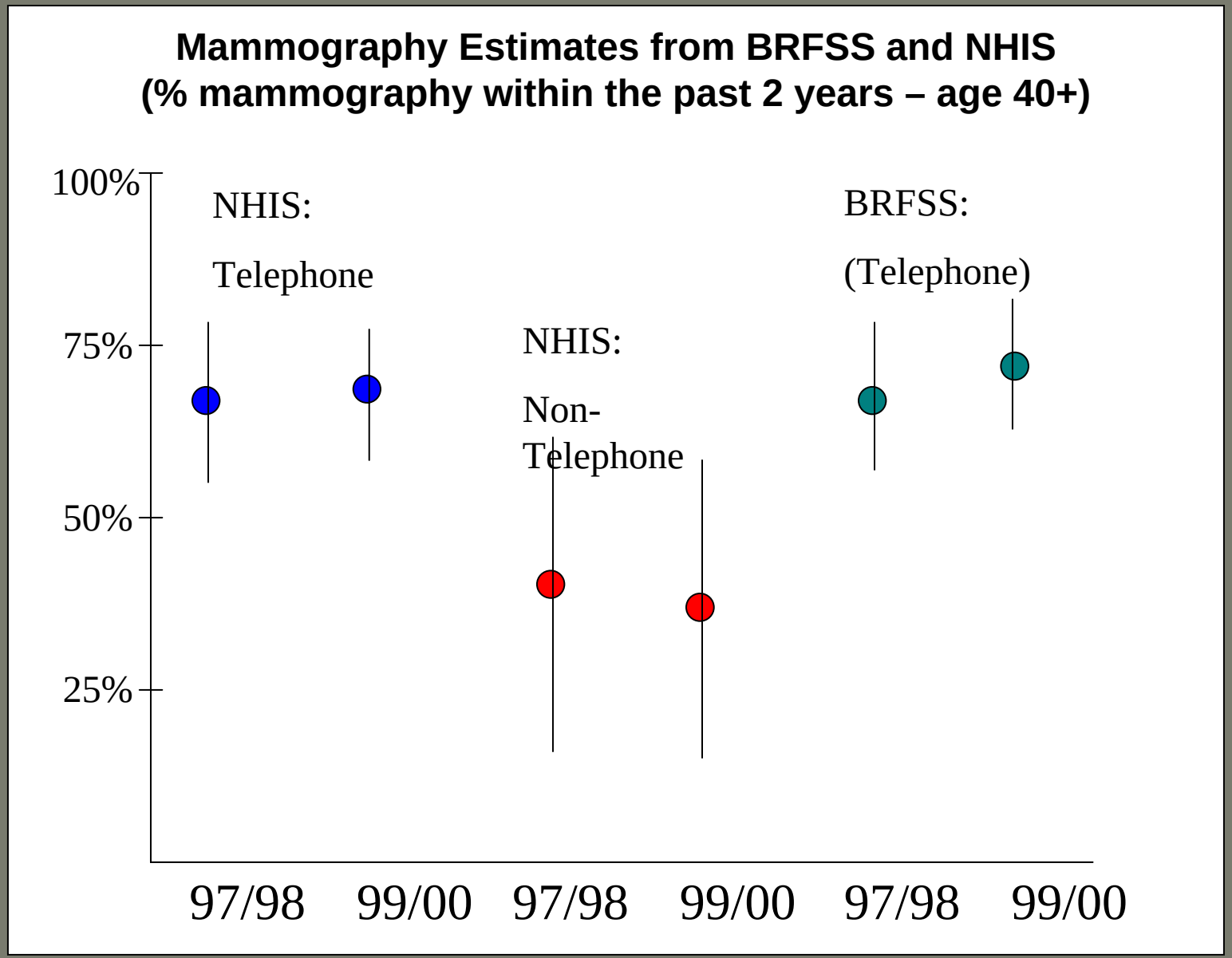
Combining Complementary Characteristics

	BRFSS	NHIS
Type	State, Telephone only	National, Face-to-face
Sample size/year	150-250 K Households	30-40 K Households
Cost/response	Low	High
Organization	CDC/States	NCHS/Census
Response rate	Lower	Higher
Coverage	Landline Telephone Residential Households, Almost all counties	Households + Sample contains about 800 counties
Available Geographical Information	State (public) County (Data on special request)	4 Regions (public) State/County (restricted access) Research Data Center

BRFSS with NHIS – Telephone Use



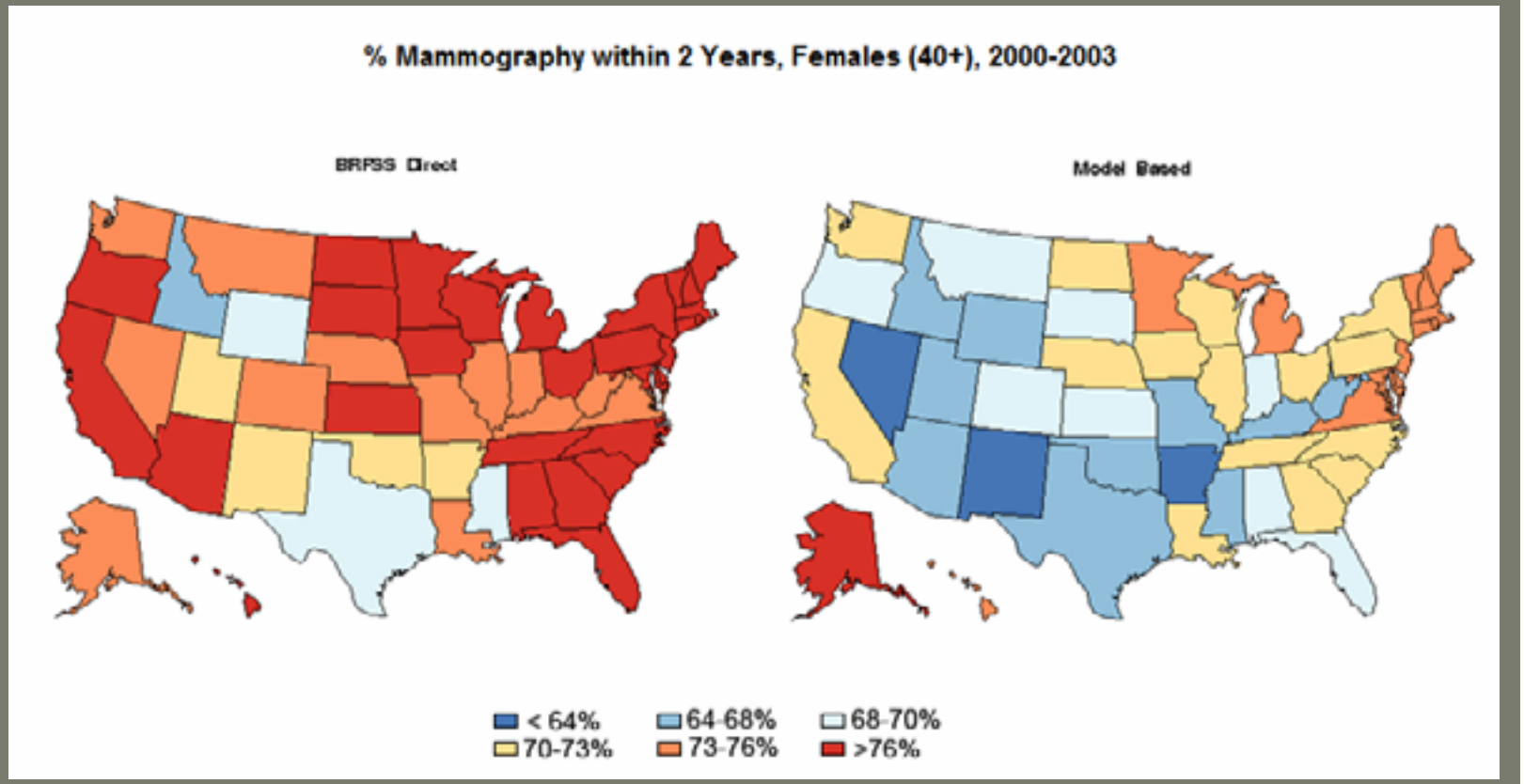
BRFSS with NHIS – Telephone Bias



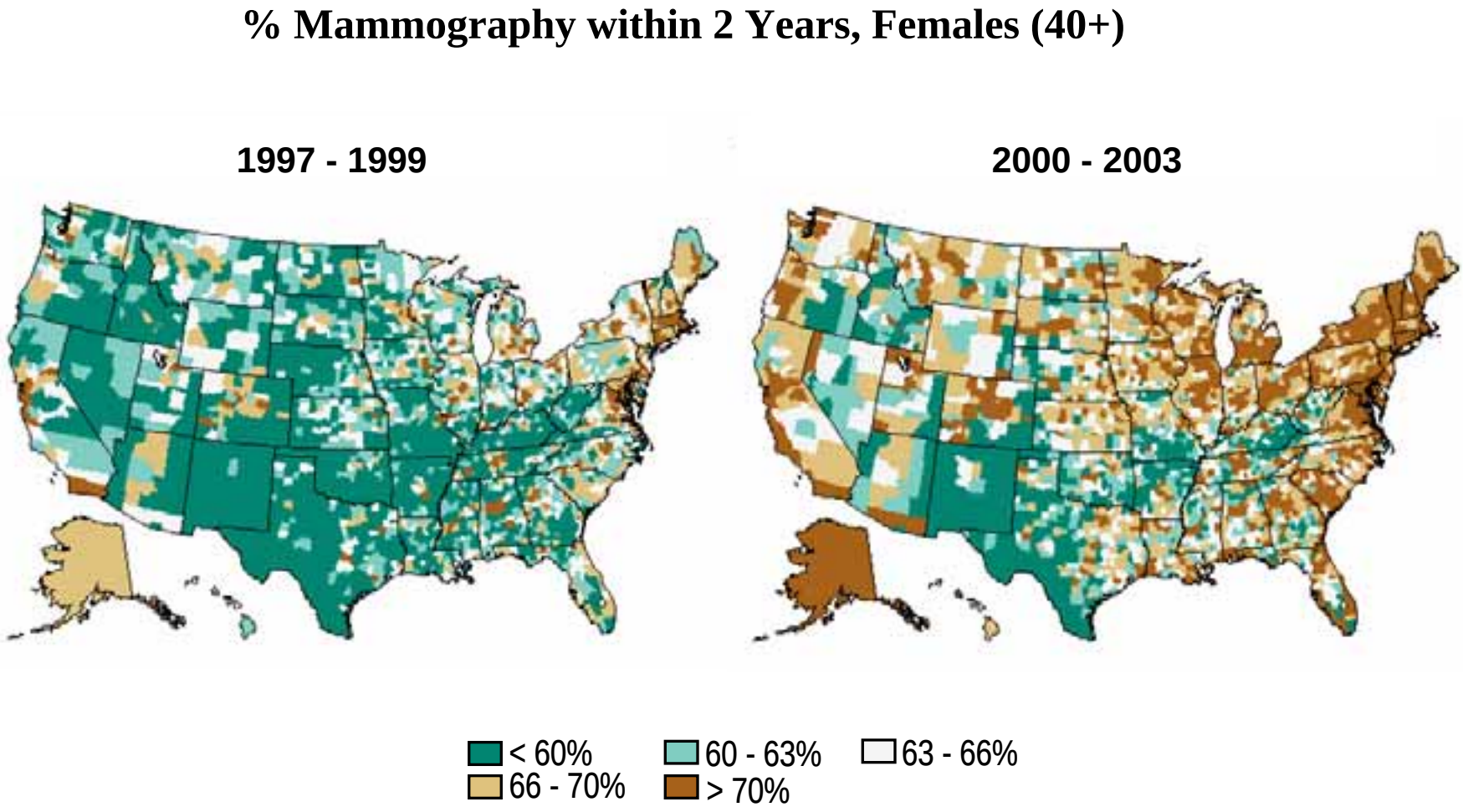
BRFSS and NHIS Model

- County-level model with:
  - NHIS response: households with phones
  - NHIS response: households without phones
  - BRFSS response
- Model covariates include:
  - Demographics
  - Socio-economic factors, crime rate
  - Population density, urban/rural, commuting

BRFSS with NHIS – example

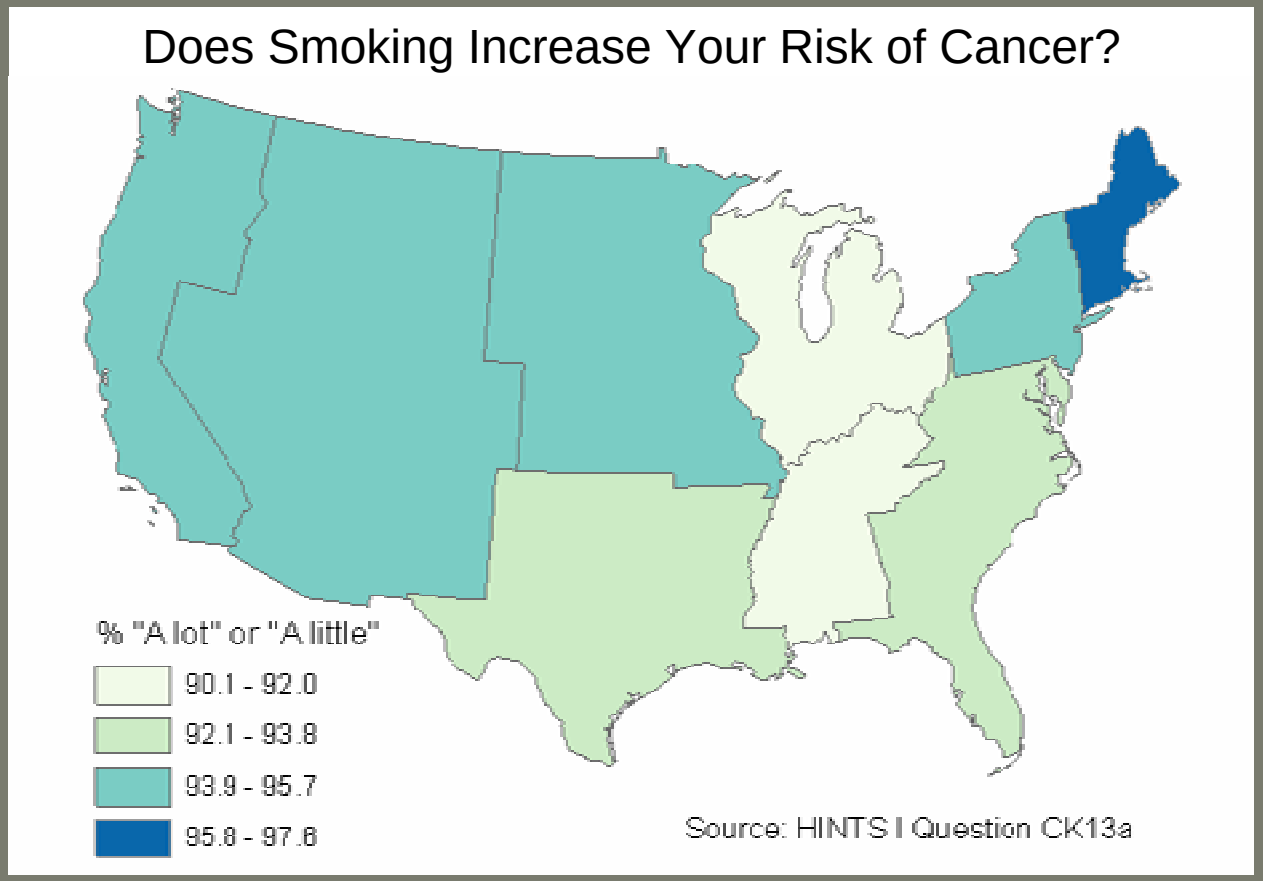


BRFSS with NHIS – example



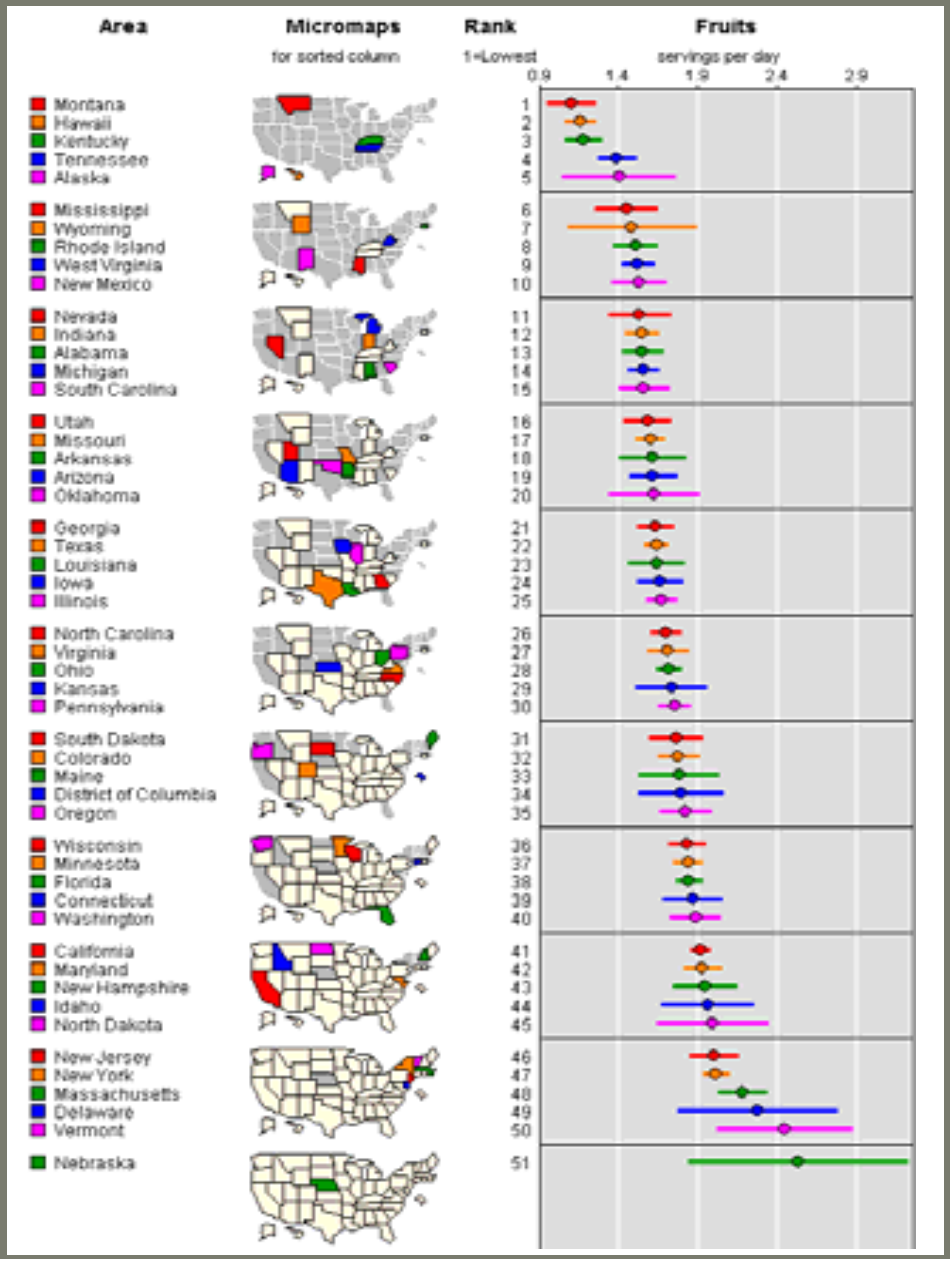
Example 3 – HINTS

- Relatively small national survey
  - About 6,000 samples from 50 states
  - Supports reporting by Census Division:



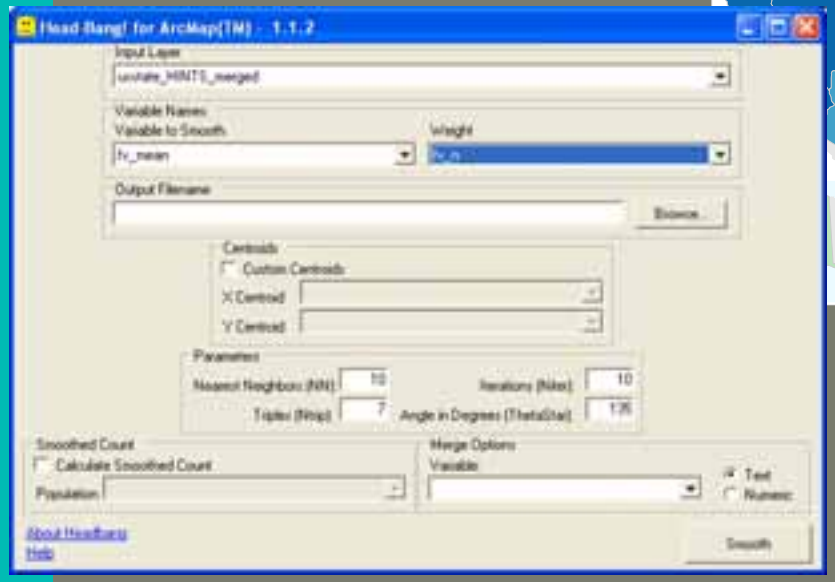
HINTS by State

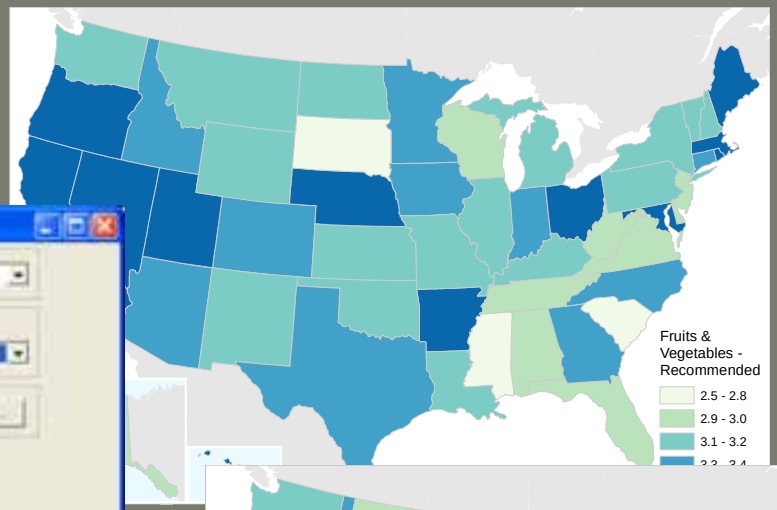
- Some states have few samples – wide confidence intervals:

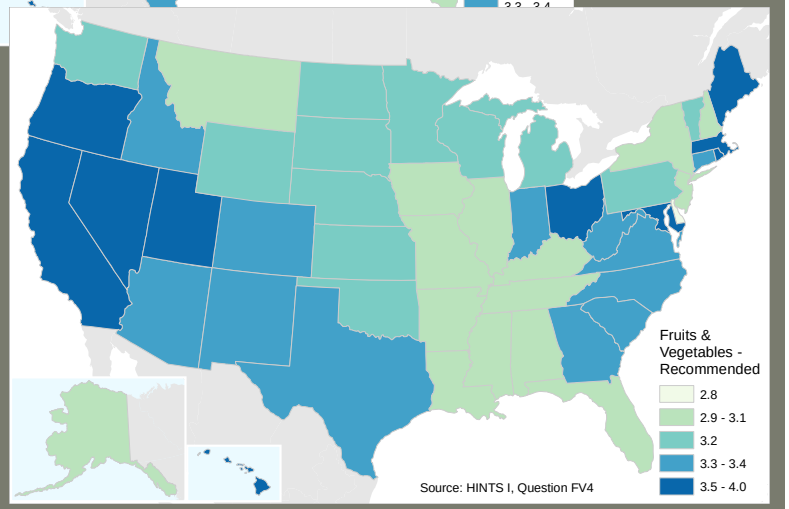


Smoothing HINTS State Data

- Headbang weighted smoothing
  - Borrows information from neighbors
  - Weighted by sample size



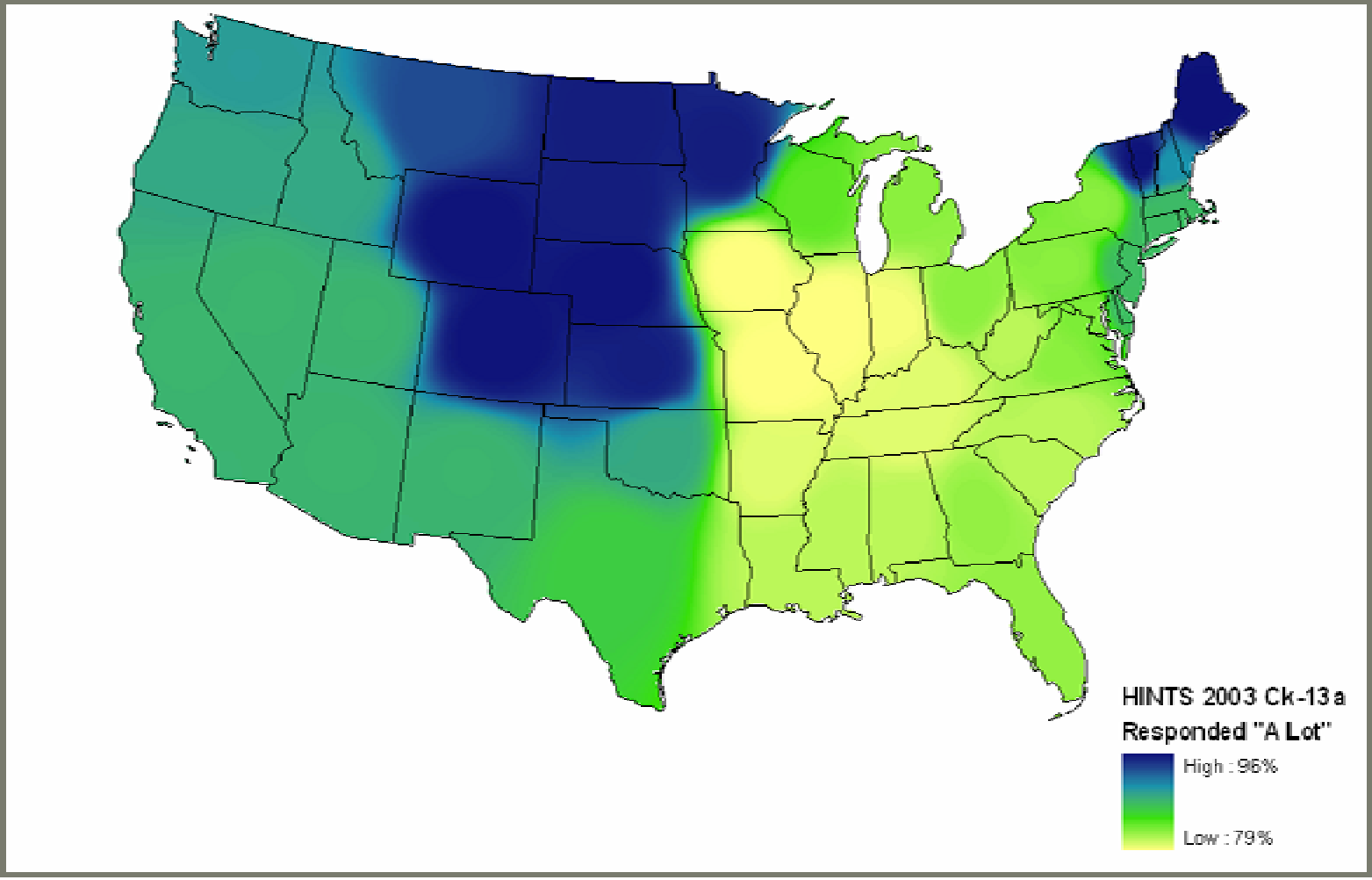




http://gis.cancer.gov/tools/headbang/

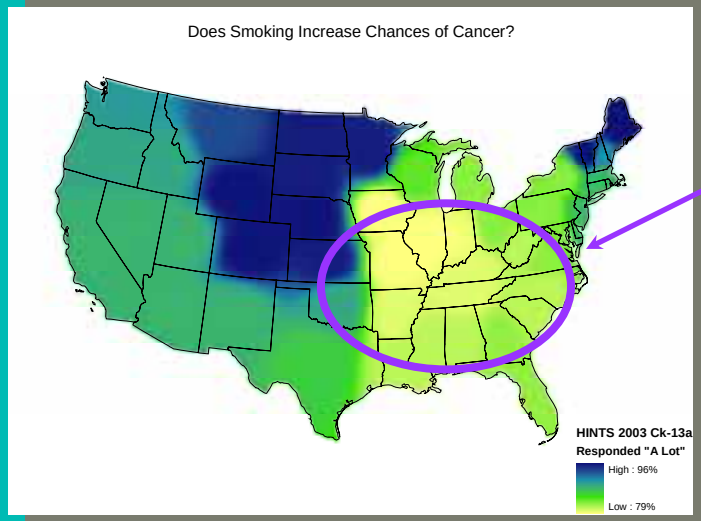
Convert to a Continuous Surface

- “Weather map” style (an isopleth map)
- Avoids transitions at state boundaries



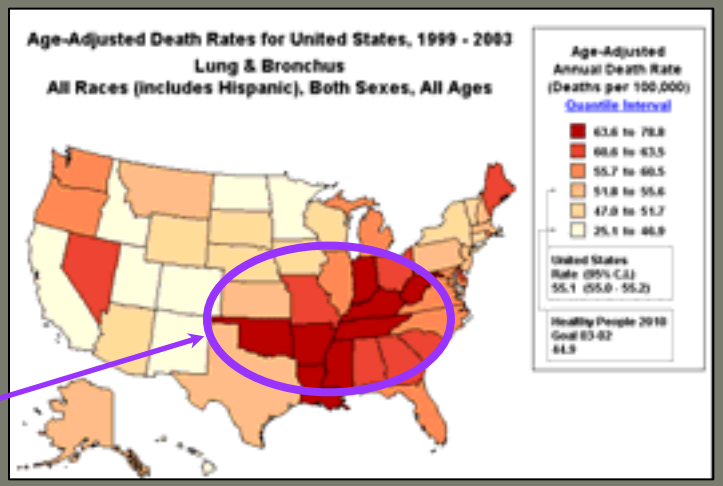
HINTS Knowledge Maps

- Maximize geographic information for communication planning
  - Not constrained by Census Divisions
  - Can show more geographic detail where there are more samples



HINTS 2003
"Smoking leads to lung cancer"

Low belief in risk of smoking



State Cancer Profiles
Lung Cancer Mortality

High lung cancer mortality rates

Conclusions

- Geographic information in survey data
  - Differences in collected information
  - It is possible to collect high quality geographic information about respondents (CHIS)
  - Can augment geographic information by combining with other data (BRFSS/NHIS)
  - Can use smoothing and isopleth mapping to maximize visualization (HINTS)
- Provide the best quality data for public health communication planning
- Can link knowledge/beliefs, behavior, and health outcomes

Thank You

Dave Stinchcomb
StinchcD@mail.nih.gov
301-594-7251