

Model-based Small Area Estimation for Cancer Screening and Smoking Related Behaviors

Benmei Liu, Ph.D.

Eric J. (Rocky) Feuer, Ph.D.

Surveillance Research Program
Division of Cancer Control and Population Science
National Cancer Institute

Presented at the NAACCR Annual Conference

June 18, 2015

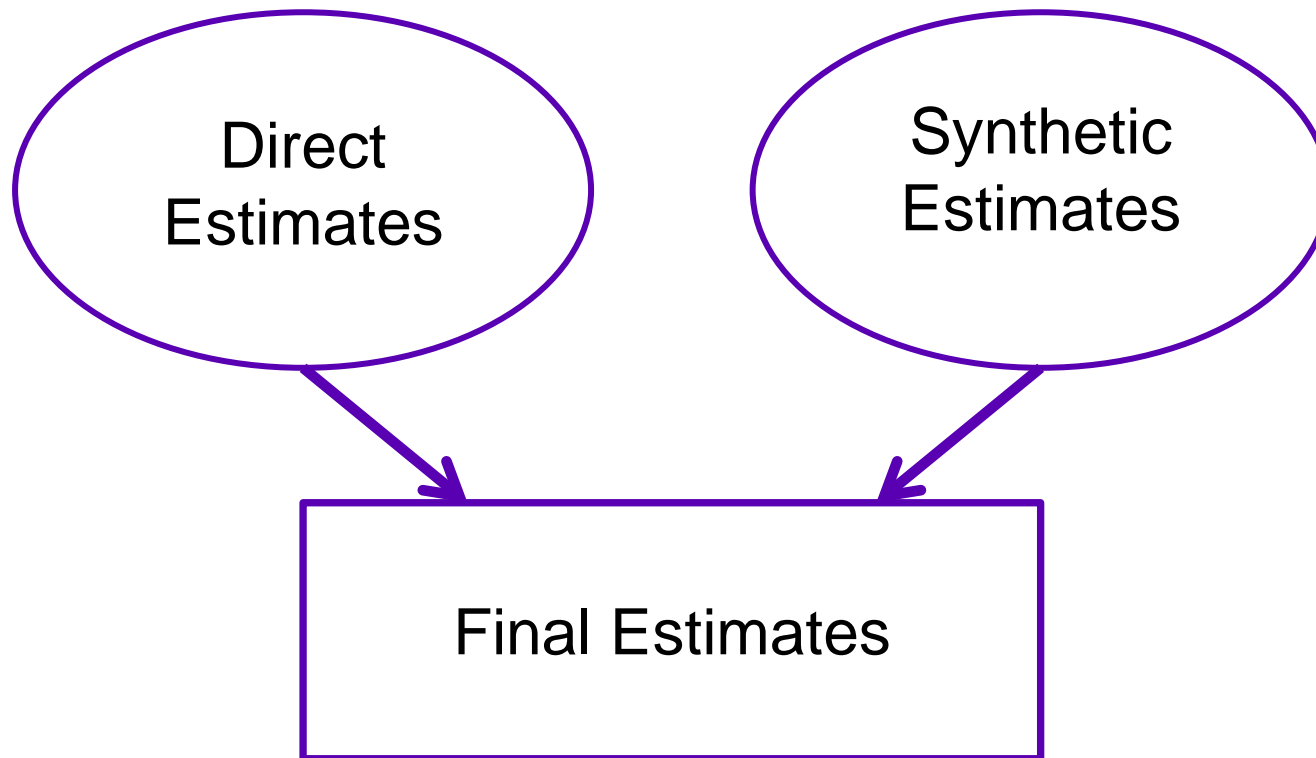
Charlotte, NC

- Background of the small area estimation (SAE) techniques
- Applications of SAE techniques at NCI
- Project I: Combining BRFSS/NHIS for cancer risk factors and screening behaviors at the state and county level
- Project II: Tobacco Related Small Area Estimation Using Data from the TUS-CPS
- Summaries

- The demand for survey estimates for small areas (small geographic areas or domains) has increased in many different areas of application (e.g., income and poverty, education, health, substance use) over the past several decades
- The standard direct estimation methods for survey data cannot provide reliable estimates due to the small sample size
- Model-based methods that combine information from multiple related sources have been developed to increase the precision

- Borrowing strength from relevant sources (Census/ Administrative information, related surveys)
- Methods of combining Information
 - Choose good small area models
 - Use good statistical methodology
- Mixed models (fixed effects + random effects) have been popularly used in the small area estimation (Rao 2003)

- The final estimates are combinations of the direct estimates and the synthetic estimates.



- Estimate cancer risk factors & screening behaviors for states and counties by combining data from Behavior Risk Factor Surveillance System (BRFSS) and National Health Interview Survey (NHIS) (<http://sae.cancer.gov/>)
- Estimate the county level prevalence of several smoking-related outcomes for the Tobacco-Use Supplement to the Current Population Survey (TUS-CPS) (Liu and Gilary 2014)
- Estimate state level proportions of people who answered 'YES' to a series of cancer-related knowledge questions using the Health Information National Trends Survey (HINTS) (http://hints.cancer.gov/docs/HINTS_IDA_Report.pdf)

- Cancer screening and risk factor data are of great interest to cancer control planners at the state and sub-state level, but accurate local statistics have been difficult to obtain
- Different surveys have different strengths
- Combining information from surveys could improve small-area estimates

Project Team

- Project led by National Cancer Institute, with collaboration by:
 - National Center for Health Statistics
 - National Center for Chronic Disease Prevention and Health Promotion
 - University of Michigan
 - University of Pennsylvania
 - Information Management Services

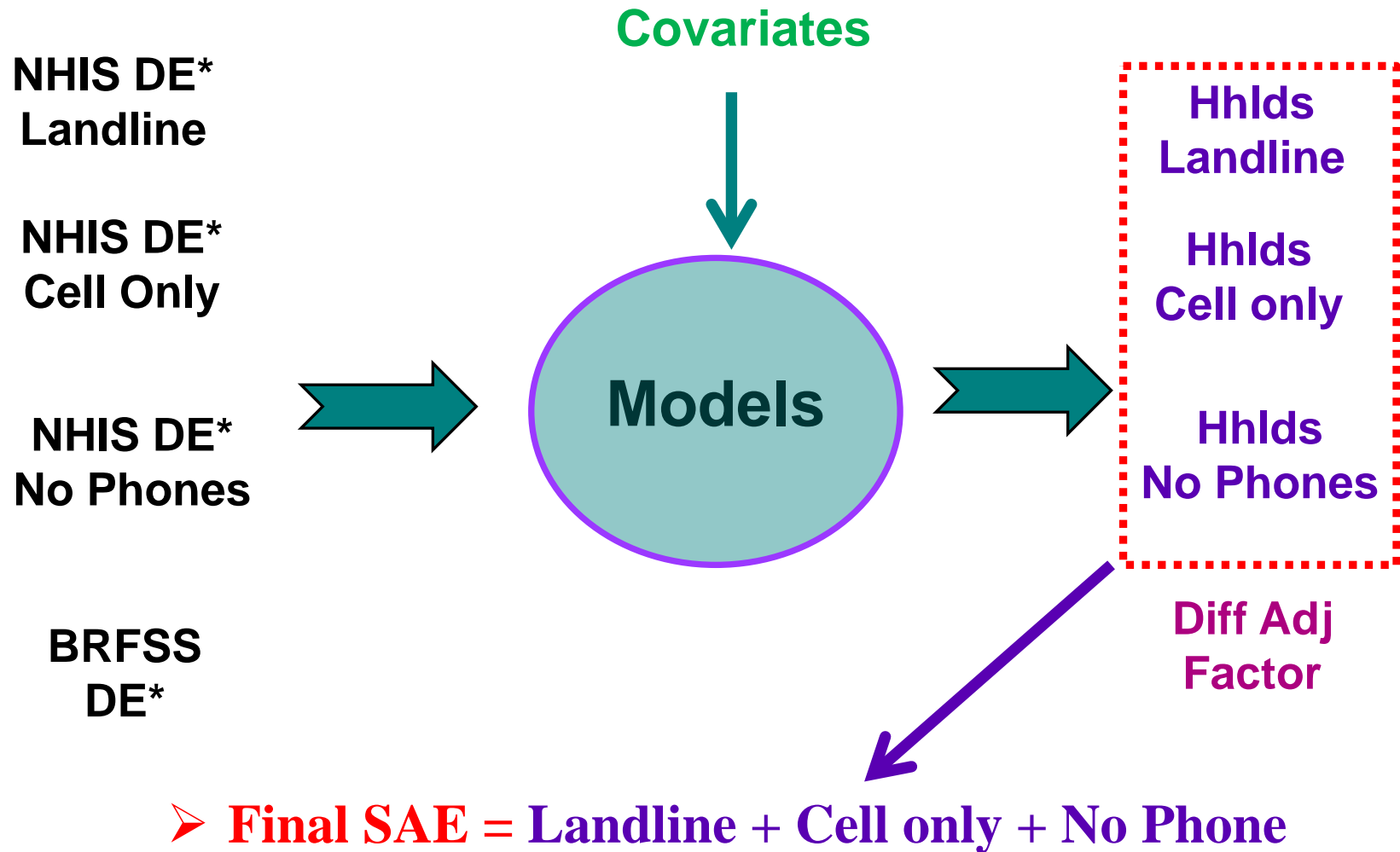
Surveys Used

- Behavioral Risk Factor Surveillance System (BRFSS) – the largest U.S. survey tracking health conditions and risk behaviors at the state and sub-state level since 1984
 - Limitations:** nonresponse bias; only hhlds with landline phones (now has a cellphone only component); noisy or no county level estimates
- National Health Interview Survey (NHIS) – the principal source of information on the health of the civilian noninstitutionalized population of U.S. since 1957
 - Limitations:** Smaller sample size; only includes data on about ¼ of U.S. counties

Can the Complementary Characteristics of NHIS be Used in Enhance BRFSS Estimates?

	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

Statistical Models and Inferences



*DE: Direct Estimate

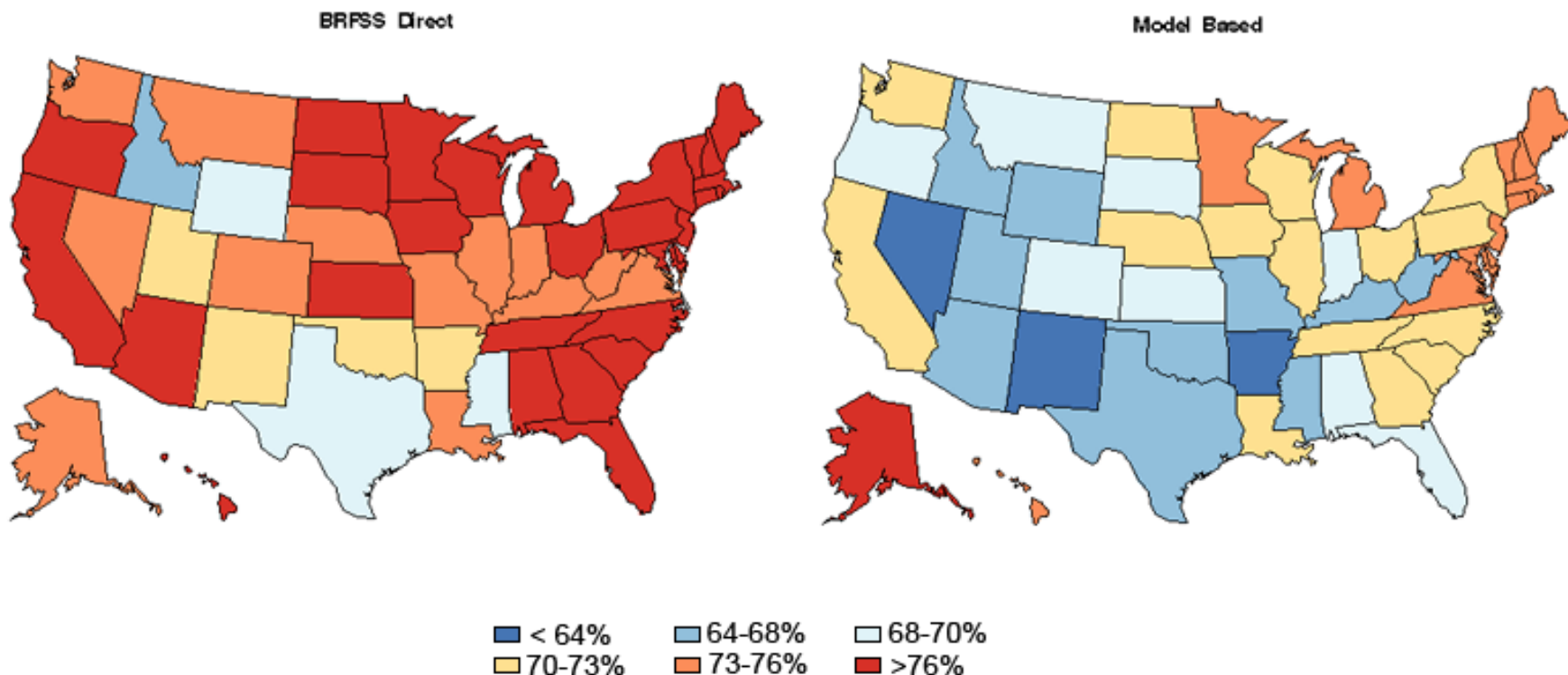
Project Results

- Bayesian methods are developed to combine information from the two surveys; also incorporated telephone coverage rates estimated from NHIS and the Census.
- Developed estimates for four time periods: 1997-1999, 2000-2003, 2004-2007, 2008-2010
 - Smoking, mammography, pap smear, Colorectal
 - Counties, health service areas, and states

Combining Surveys for Improved State Level Estimates

Supplementing BRFSS with NHIS to obtain improved estimates

% Mammography within 2 Years, Females (40+), 2000-2003



Project Dissemination

➤ Publications:

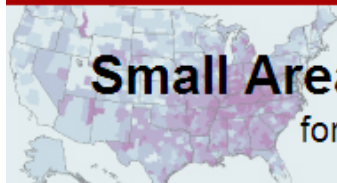
- Methods paper: Journal of the American Statistical Association 2007
- Application paper: Public health Reports 2010
- Two+ new papers are planned to write

➤ Websites:

sae.cancer.gov

<http://statecancerprofiles.cancer.gov/>

<http://seer.cancer.gov/seerstat/variables/countyattribs/>



Small Area Estimates

for Cancer Risk Factors & Screening Behaviors

Search

Home

Understanding this site

Small Area Estimation Overview

Methodology

Data Sources

Limitations & Uses

Rationale

Available Estimates ▼

Resources ▼

Frequently Asked Questions

Contact Us

Small Area Estimates for States, Counties, & Health Service Areas

These estimates use two surveys:

- Behavioral Risk Factor Surveillance System (BRFSS)
- National Health Interview Survey (NHIS)

The two surveys are combined using novel statistical methodology. This yields estimates for states that strengthen BRFSS results with the higher response rate and coverage of non-telephone households available from NHIS and bias-corrected estimates for counties by utilizing data from both surveys and information from other areas with similar characteristics.

Small area estimates are available for:

- [Current Smoking Prevalence \(Age 18+\)](#)
- [Ever Smoking Prevalence \(Age 18+\)](#)
- [Mammography Prevalence within 2 Years \(Age 40+\)](#)
- [Pap Smear Test Prevalence within 3 Years \(Age 18+\)](#)

Available estimates are for the time periods 1997-1999 and 2000-2003. Years were grouped to provide reasonable sample sizes in each area to inform our models.

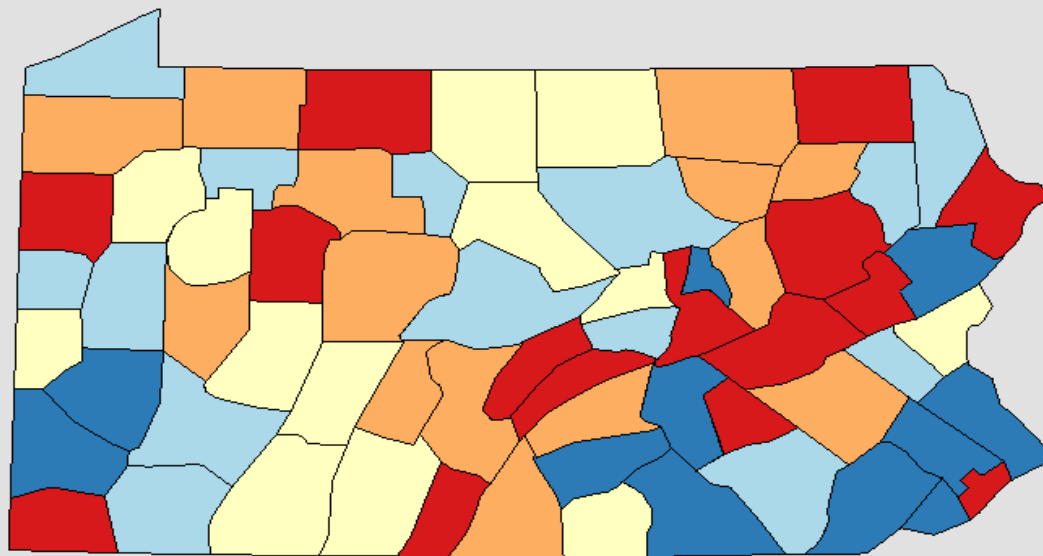
Model-based estimates incorporate data from other areas with similar profiles, especially where sample sizes are small. See [Limitations & Uses](#).

The model-based estimates produced and presented on this website are not directly estimated BRFSS or NHIS area level estimates, although BRFSS and NHIS data was incorporated into the models.








2007 SPAIG Award Winner – The American Statistical Association presented this award to the collaborating organizations responsible for developing the methodology for these estimates.

**Screening and Risk Factors for Pennsylvania
(2008-2010 County Level Modeled Estimate Combining BRFSS & NHIS)
Ever Had Colorectal Endoscopy (Sigmoidoscopy or Colonoscopy)
All Races (includes Hispanic), Both Sexes, Ages 50+**



Ever Had Colorectal
Endoscopy
(Sigmoidoscopy or
Colonoscopy)
(Percent of Respondents)

[Quantile Interval](#)

-  62.8 to 70.2
-  60.0 to 62.8
-  57.0 to 60.0
-  54.4 to 57.0
-  47.1 to 54.4

Notes:

Created by statecancerprofiles.cancer.gov on 06/12/2015 10:20 am.

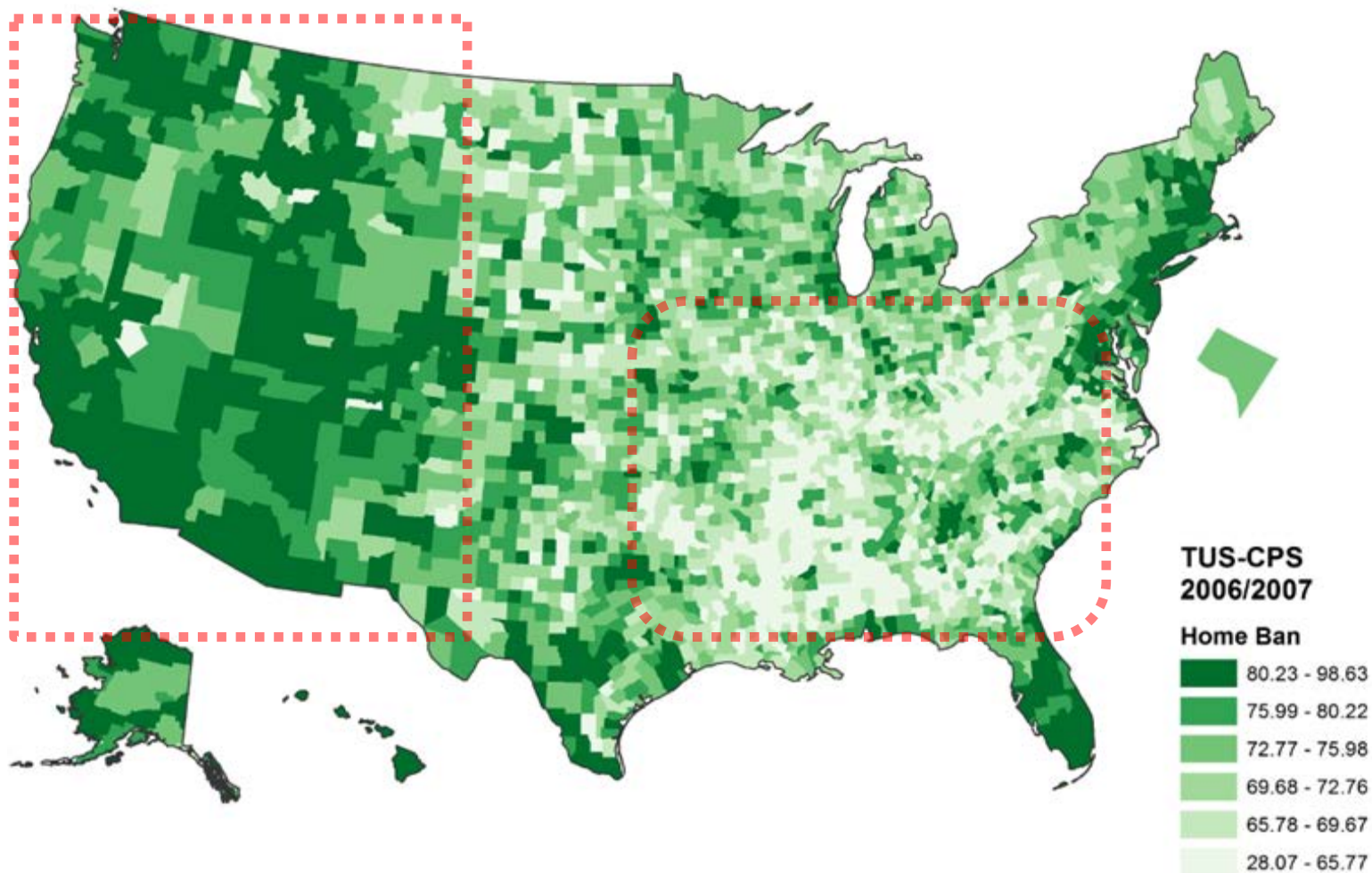
Small Area Estimates is the source for this data.

Estimates are based on a statistical model which combines information from the Behavioral Risk Factor Surveillance System and the National Health Interview Survey to correct for nonresponse and undercoverage bias and are enhanced in small areas by borrowing information from similar areas across the nation. For more information, visit <http://sae.cancer.gov/>.

Project II: Tobacco Related Small Area Estimation Using Data from the TUS-CPS

- Model-based county level estimates for the following key measures (for population aged 18+):
 1. Percent of population currently smoking
 2. Percent of population that has ever smoked
 3. Percent of population that has quit for 24+ hours, among those who have smoked within the past year
 4. Percent of population governed by a smoke-free workplace policy
 - (Workplace has an official smoking policy: Smoking Not allowed in ANY public areas and work areas)
 5. Percent of population governed by a smoke-free home rule
 - (No one is allowed to smoke anywhere INSIDE YOUR HOME)
- Collaboration among NCI, the Census Bureau, and the University of Maryland
- ✓ Important to the Tobacco Control Research Branch

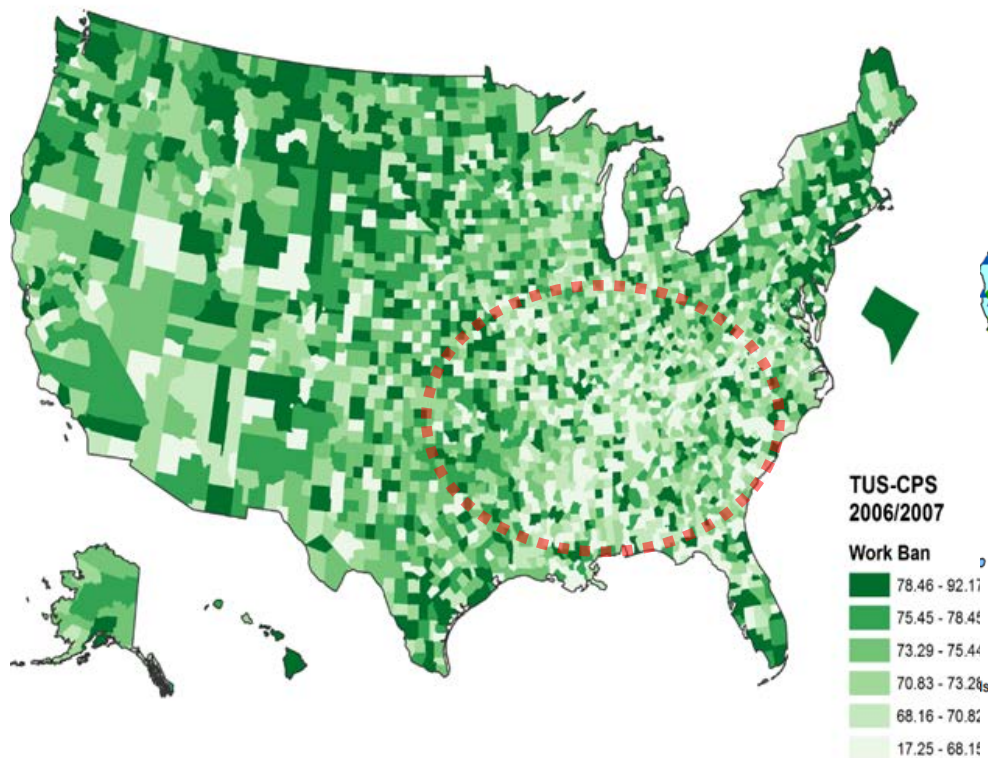
Model-based Estimates for Percent of Population Governed by a Smoke-free Home Rule* Among Age 18+ : TUS-CPS 06/07



*No one is allowed to smoke anywhere **INSIDE**
YOUR HOME

Model-based Estimates for Percent of Population Governed by a Smoke-free Workplace Policy* Among age 18+: TUS-CPS 06/07

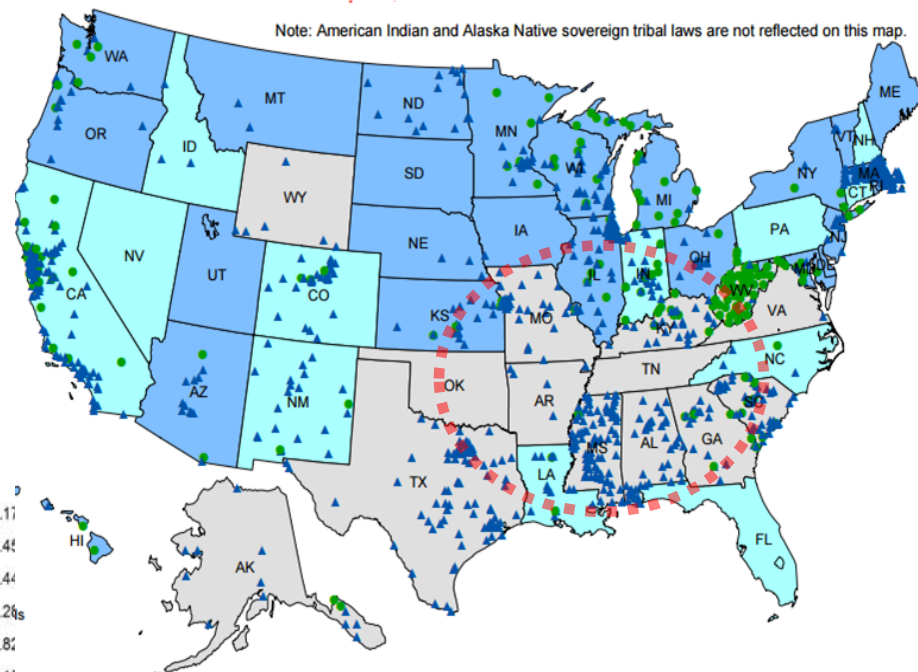
Individual Self-Reported



Law Legislations

United States 100% Smokefree Air Laws (as of April 2, 2015)

Note: American Indian and Alaska Native sovereign tribal laws are not reflected on this map.



<http://www.no-smoke.org/pdf/100Map.pdf>

*Workplace has an official smoking policy: Smoking Not allowed in ANY public areas and work areas

Summary

- The model-based SAE techniques represent an effective means of generating estimates where there is small (or zero) state or county sample.
- The SAE results, which are released and disseminated at several NCI's websites provide a useful resource for the broad cancer surveillance society to fulfill multiple needs.
- We are working on to update the websites to include estimates for the two recent data periods from project 1 and to add estimates from project II.

Any Questions?

Thank you!

Contact info:

Benmei Liu

liub2@mail.nih.gov