Overview & Background of TUS-CPS

Carolyn Reyes-Guzman, PhD, MPH

Understanding Tobacco Use Trends: Leveraging Harmonized Data from the U.S. Tobacco Use Supplement – Current Population Survey, 1992-2015

February 20, 2019

NIH > NATIONAL CANCER INSTITUTE

Tobacco Use Supplement
Current Population Survey

Current Population Survey

- The CPS is a monthly labor force survey conducted in about 56,000 interviewed households across the U.S.

- Each household is interviewed monthly for 4 consecutive months one year, and again for the corresponding time period a year later.

  - Can obtain reliable month-to-month and year-to-year comparisons

- Secondary purpose of CPS is collecting information on demographic characteristics (age, sex, race, marital status, educational attainment, family relationship, occupation, and industry)

  - Additional questions included as Supplements on health, education, income, and previous work experience

NATIONAL CANCER INSTITUTE

TUS-CPS Background

- NCI primarily sponsored TUS-CPS as part of the US Census Bureau's Current Population Survey every 3-4 years since 1992-93

- FDA currently a co-sponsor since 2014-15

- CDC co-sponsored from 2001-02 through 2006-07

- Most recent publicly released data are from the 2014-2015 TUS-CPS wave; 2018-2019 wave currently being fielded.

- Previous waves are 1992-1993, 1995-1996, 1998-1999, 2000, 2001-2002, 2003, 2006-2007, and 2010-2011

  - Data collected during 3 time points for every cycle, except 2000 (2 points)

TUS-CPS Longitudinal Cohort

- NCI additionally sponsored a special longitudinal TUS-CPS cohort administered in both May 2010 and May 2011

- Part of 2001-2002 wave (panels 1-3 of February 2002) overlap with part of 2003 wave (panels 5-7 of February 2003), providing another opportunity to analyze longitudinal data

- Overlap between January and May 1999 and January and May 2000, respectively

What is TUS useful for?

- TUS data can be used by researchers to:

    - Monitor tobacco control progress and assess long-term cross-sectional population trends

    - Track topics such as tobacco health disparities, which may require larger sample sizes

    - Evaluate tobacco control programs

    - Examine national, state and county*-level data

- Data from the TUS can be linked to other CPS supplements

*where available

Recently-Added TUS Topics

- More recent series have included more detailed information on:

    - emerging tobacco products

    - use of flavored non-cigarette tobacco products

    - attitudes towards multi-unit housing smoke-free policies (since 2014-15)

        - These questions not included in current harmonized data file, but will be incorporated after 2018-19 data becomes available

New Questions in 2018-2019 cycle

- Expanded non-cigarette products flavor to get at **flavor categories**, including for e-cigarette "no flavor" vs "tobacco or menthol"

- **Device type** for e-cigarettes

- **E-cigarette Price** in addition to cigarette price (existing)

- Identify **18-34 yr.** olds who **haven't smoked 100+** cigarettes **but do currently smoke** cigarettes

- Added back **some treatment aid** items that were previously dropped

- *"Vape-free"* **home** rules and "*vape-free*" **workplace** policies

- *Awareness* item for the court-mandated *TI corrective statements* to be placed in newspapers & disseminated through other vehicles.

Trends in Tobacco Use Prevalence and Attitudes: Illustrations from the TUS-CPS

Gordon Willis, PhD

February 2019

NIH NATIONAL CANCER INSTITUTE

Tobacco Use Supplement Current Population Survey

Purposes for assessing trends - within a repeated, cross-sectional tobacco survey

1) Descriptive: Assessing progress in an area of public health

   - Reaching Health People Year 20X0 Objectives

2) Predictive: "Trend Analysis" attempts to project future behavior

   - Predicting when the US will achieve <10% adult cigarette smoking

3) Differential: Showing how a trend over time varies for a particular class or demographic group

   - Reduction in smoking by race/ethnicity

4) Inferential: Relating discontinuities (jumps) to critical events

   - Increase in high school e-cig use after introduction of JUUL

Figure 1. Tobacco Use Supplement to the Current Population Survey: Total U.S. and Regional Trends in Current (Every Day or Some Day) Cigarette Smoking Prevalence, 1992–2015
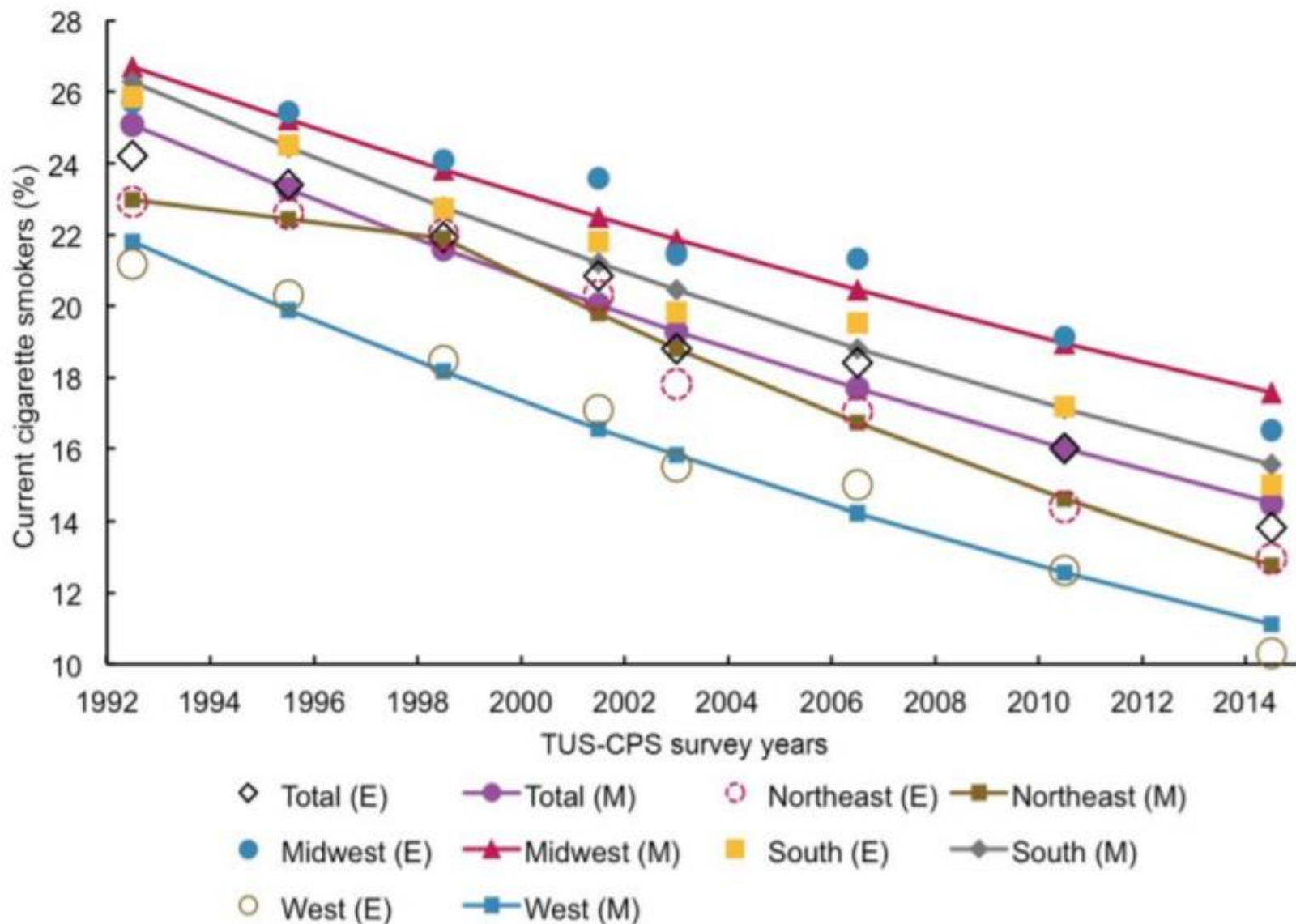

Legend:
- ◇ Total (E)
- ●— Total (M)
- ◌ Northeast (E)
- ■— Northeast (M)
- ● Midwest (E)
- ▲— Midwest (M)
- ■ South (E)
- ◆— South (M)
- ○ West (E)
- ■— West (M)

Axis labels: Current cigarette smokers (%); TUS-CPS survey years

Figure 3. Trends in Percentage of Current Smokers Reporting Menthol as Usual Type of Cigarette, Total United States, by Sex, TUS-CPS 2003–2015
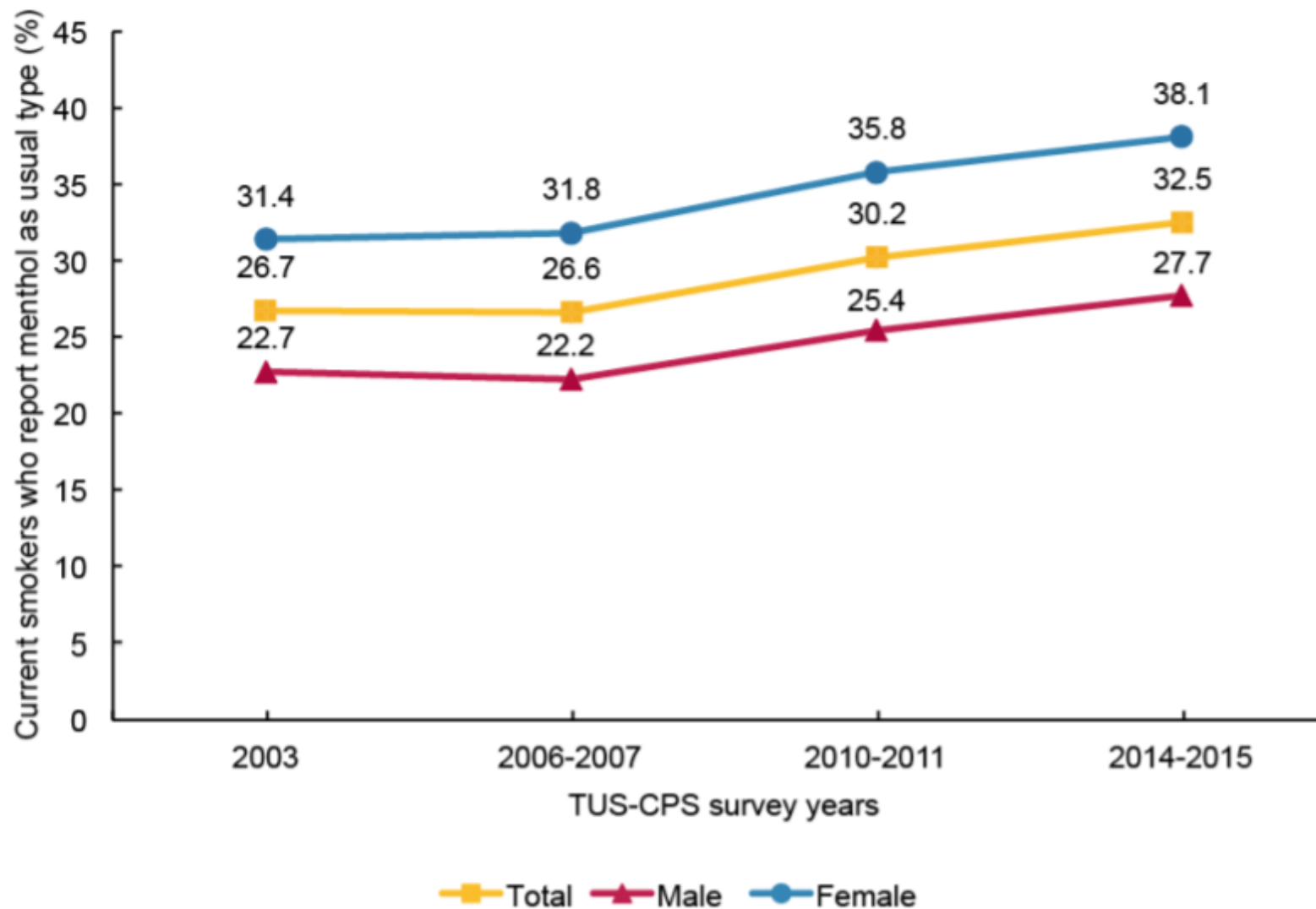
Figure 4. Trends in Percentage of Current Smokers Reporting Menthol as Usual Type of Cigarette, by Race/Ethnicity, TUS-CPS 2003–2015
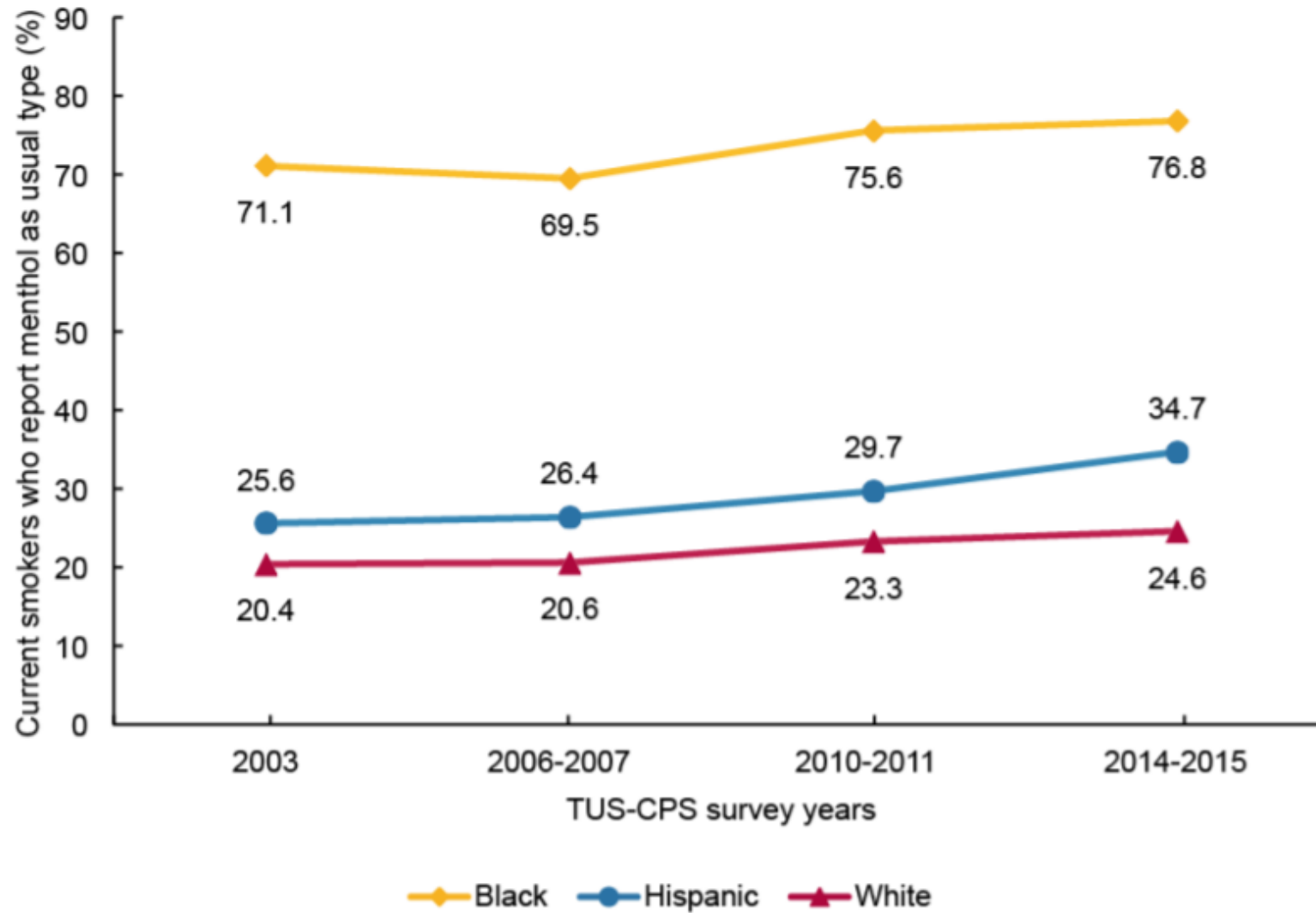
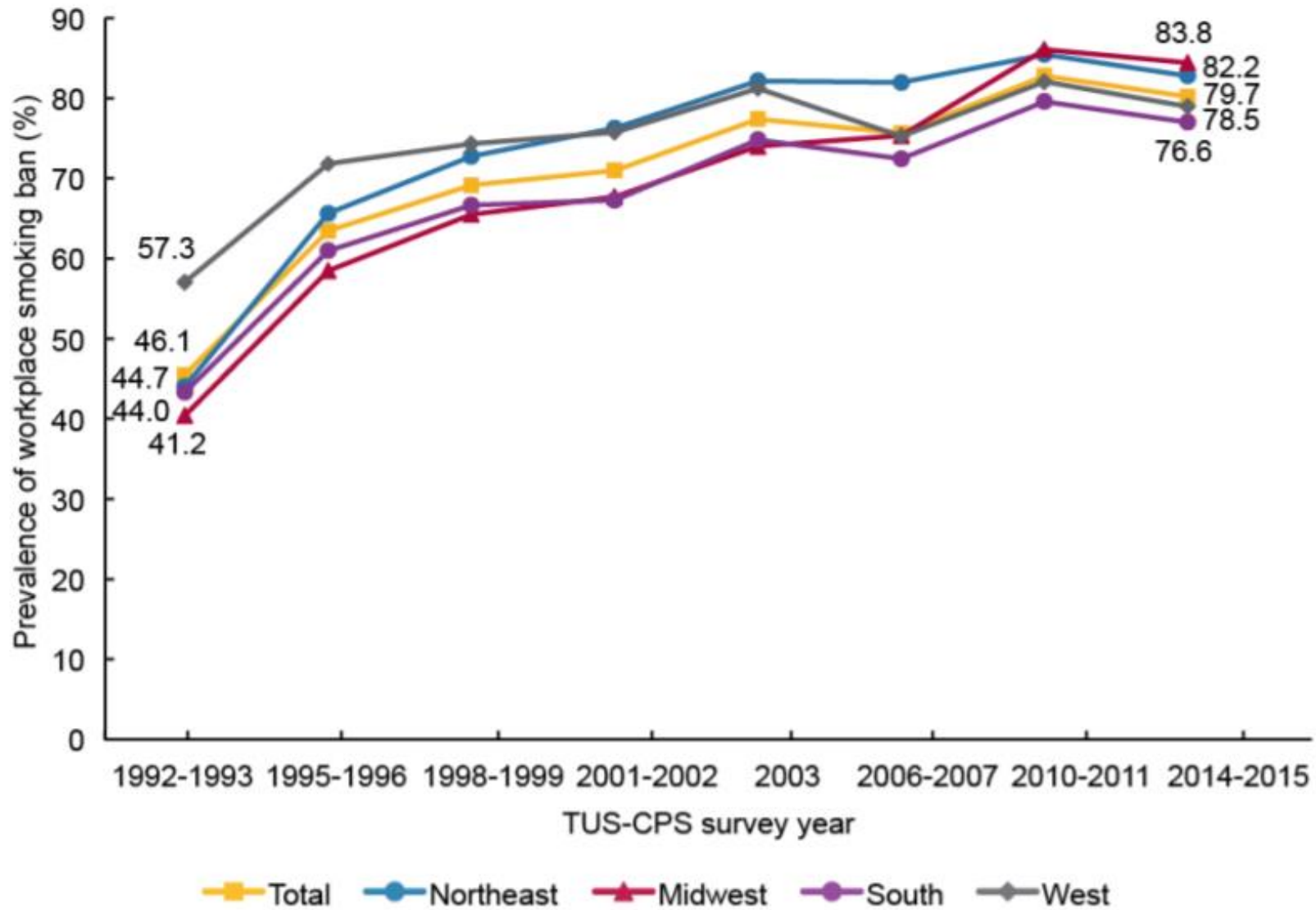Figure 6. Trends in Workplace Smoking Ban, by Region, TUS-CPS 1992–2015

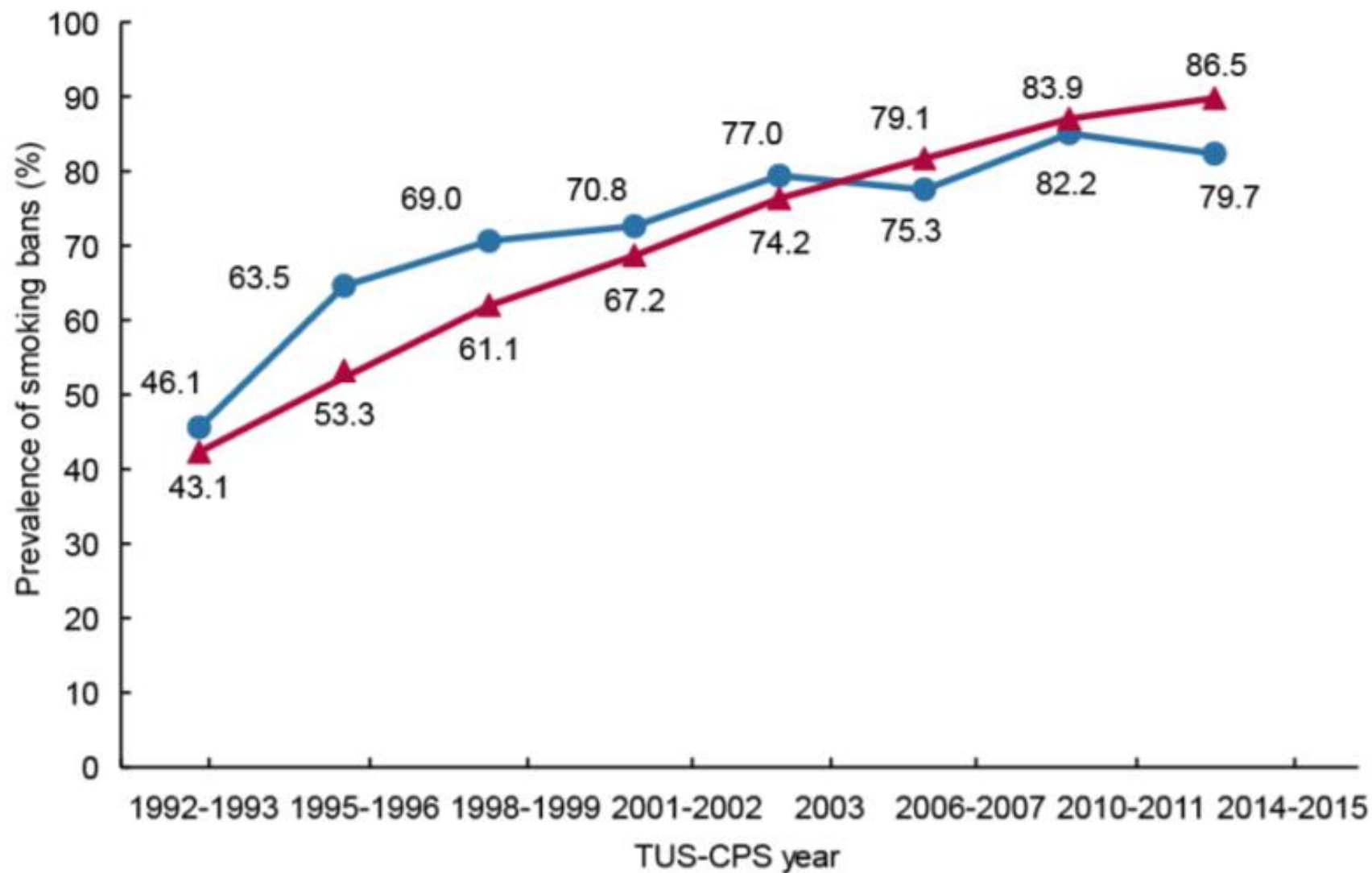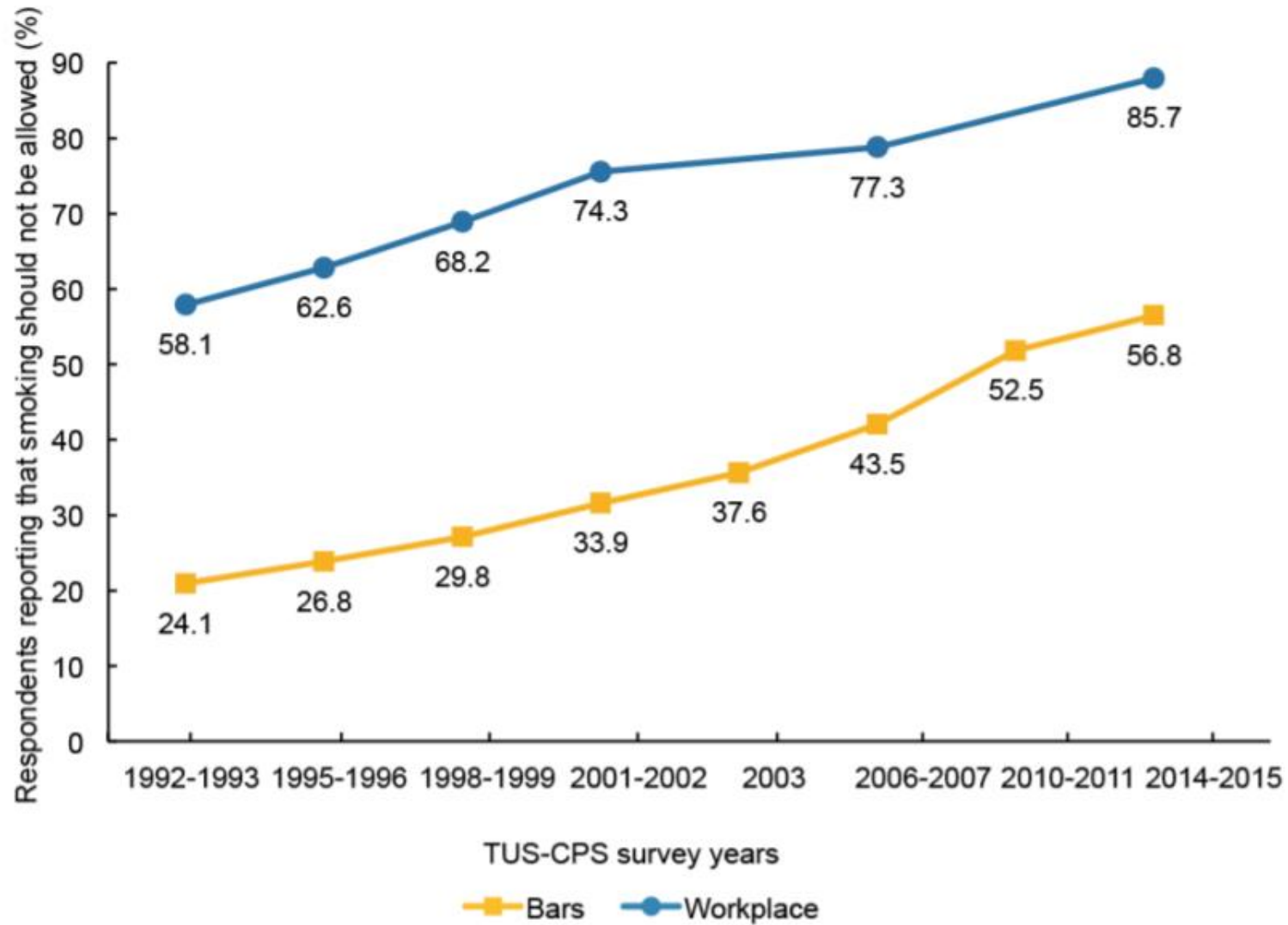Figure 7. Home and Workplace Smoking Bans Across TUS-CPS Waves, 1992–2015

**Figure 8. Trends in Attitudes Toward Smoking Bans
In Indoor Workplaces and in Bars/Cocktail Lounges, TUS-CPS 1992–2015**

Overview

Unique Aspects: Stems from the "3 L's"

- **LARGE** sample SIZE, **Longevity,** and **Linkages** are KEY to many of the *Unique* features along with *Unique* **Content** and **Design**

  - Derived from the CPS content

  - Examine detailed tobacco health disparities

  - Study by *small* / *novel* units of geography

  - CPS design provides opportunity for longitudinal/cohort evaluation

  - CPS Linkages through NLMS & other CPS supplements

- **Unique** features' applications often combine > 1 of these 5 aspects

Unique Features

Uniqueness: Derived from the CPS Content

- Deep Dive on Occupation & Industry
  - bartenders, construction workers, nurses subcategories (LPNs, RNs, etc.)
  - food service workers by job category & type of venue (food prep, dishwashers, waitresses/waiters, servers, hosts/ restaurants, cafeterias)
- Work Force/ Employment Considerations

  - **Congressional Budget Office** used TUS to assess impact of raising the Federal tobacco excise tax - effect of smoking on earnings perspective
  - use of "# of sick days" as a health outcome in tobacco control studies
  - immigration and access to fringe benefits by Maclean et. al., 2018
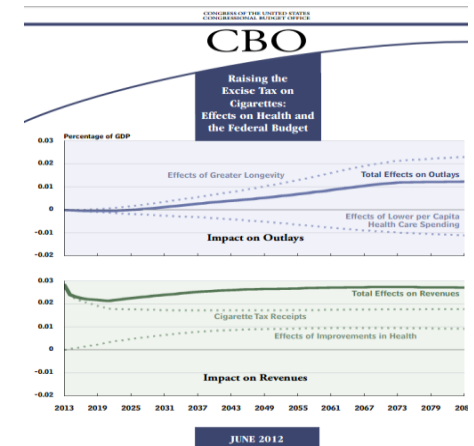
Uniqueness: Detailed Tobacco Health Disparities

- Racial/ethnic differences

  - Hispanics by country of origin and geography (e.g., Hispanics in NYC)

  - South Asians (India, Pakistan, Sri Lanka, Bangladesh)

  - intergenerational differences between blacks by country of origin

- Rural/urban differences

  - 10 HHS regions across the U.S.

  - restricted data to study rural MS Delta & Appalachia with a new isolation index

- Immigration & acculturation

- Other socio-economic differences

  - projected impacts of cigarette floor **price** & **excise tax policies** (Golden et. al., 2016)
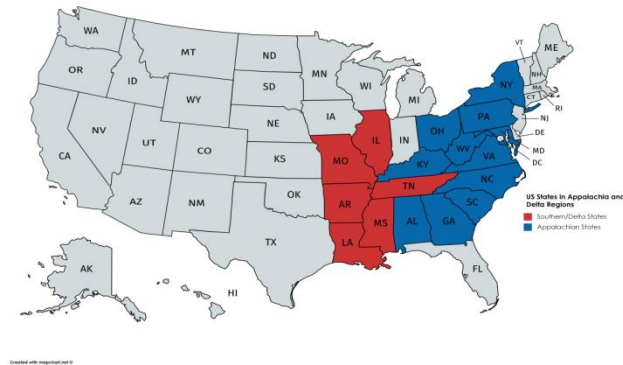
Uniqueness: Small/novel Geographic units

- State and sub-state levels, e.g., major cities are in the sample with certainty

- Major media markets identified which may cross state boundaries

  - Emery et. al., 2012 linked smoking related television advertising ratings to TUS data

  - Niederdeppe et. al., 2018 market level exposure to state anti-smoking campaigns and public support for tobacco control policy

- Comparison of states/ groups of states by tobacco control policy, attitudes, social norms impact on tobacco use/cessation treatment

  - Pierce et. al., 2018- changes in tobacco control policy, norms in CA vs. rest of US- impact on long-term trends in per capita cigarette consumption

  - Dahne et. al., 2017 –studied state tobacco control policies differential impact on cessation treatment utilization across established tobacco disparities groups

Uniqueness: Longitudinal/cohort evaluation from CPS design

- CPS panel design allowed us to efficiently conduct 3 short-term cohorts' evaluations

  - Vijayaraghavan et. al., 2018 – studied income disparities and diffusion of smoke-free homes impact on smoking cessation using two cohorts

  - Chang et. al., 2018, Zhu et. al., 2009– transitions between cigarette and smoke-less tobacco use from 2002-03 and 2010-11, and 2002-03 cohort data, respectively

  - Leas et. al., 2018 – effectiveness of pharmaceutical smoking cessation aids in a cohort of smokers 2010-11 and 2002-03

  - Nagelhout et. al., 2015 – smokers support smoke-free laws to help them quit smoking, 1999-2000 cohort analysis

- Reliability of self-reported smoking history measures

  - Soulakova et. al., 2012-2018

Uniqueness: CPS Linkages through CPS Supplements

- Special direct linkages to disease outcomes data with National Longitudinal Mortality Study (NLMS)

  - Linkages to the CPS Annual Social and Economic Supplement (March ASEC) through the CPS panel design

  - CPS data linked to National Death Index (through 2011), NCI SEER Cancer Registry, and CMS Medicare
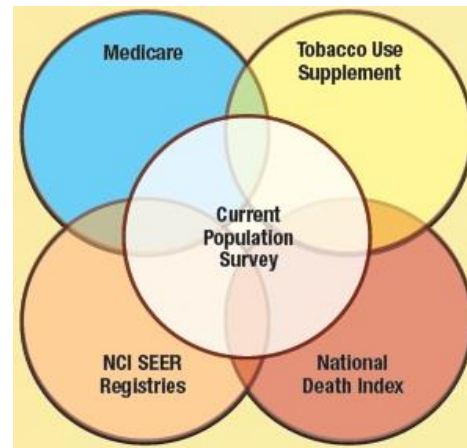
Uniqueness: CPS Linkages through CPS Supplements

- Linkage to ASEC providing detailed economic, occupational, social, and health insurance data

  - examine effect of Medicaid coverage of tobacco dependence treatments on quitting attempts and intention to quit by Medicaid smokers (Liu et. al, 2010)

  - attitudes towards multi-unit housing smoke-free policies, tobacco use and type of housing

- Linkage to the American Time Use Survey (ATUS) sponsored by BLS

  - provides time estimates that Americans spend in various primary activities for a given day of the week based on a 24 hr. activity diary

  - sample drawn from CPS samples 1 month after they have completed their 8 mo. panel participation

  - Song, 2012- pattern of daily activities for smokers, former smokers, and never smokers

Uniqueness: CPS Linkages through CPS Supplements

- Voting and Registration can provide data about the ability to influence policy

  - Hersch et. al. , 2004 – compare influence of TUS attitudes to smoking in public places with state smoke-free laws and participation in voting

- Food Insecurity

  - Farrelly et. al., 2017 – Trends in food insecurity and cigarette smoking 1998-2011

- Veterans – info about branch, length, and period of military service

- Internet and Computer Use

  - R21 (Stillman & Wewers described earlier) Supplement (Stillman & Emery)- merging info from this to TUS to develop feasible future interventions in the MS Delta & Appalachia rural/urban areas

Linkages to CPS Supplements

- March Annual Social and Economic (ASEC) Supplement: https://www.census.gov/prod/techdoc/cps/cpsmar03.pdf

- Voting and Registration: https://www.census.gov/topics/public-sector/voting/data.html

- Computer and Internet Use: https://www.census.gov/topics/population/computer-internet.html

- Food Security: https://www.ers.usda.gov/data-products/food-security-in-the-united-states/

- American Time Use Supplement (ATUS): https://www.bls.gov/tus/

- National Longitudinal Mortality Study: https://www.census.gov/did/www/nlms/index.html

Small Area Estimation

RATIONALE: Why? & What?

- TUS-CPS is designed to produce reliable estimates at national, state, & some sub-state-levels

- Policy makers, cancer control planners and researchers often need county level data for tobacco measures

- *County level* estimates (n=3,137) for the following key measures (2010/2011 TUS, age 18+):
  - % of pop. currently smoking
  - % of pop. that have ever smoked
  - % of pop. quit for 24+ hours within the **past year**, among those **currently** smoking at **some time** during the **past year**
  - % of pop. **reporting** a smoke-free workplace **policy**
  - % of pop. **reporting** a smoke-free home **rule**

Small Area Estimates: How?

- Collaboration among NCI, Census Bureau and U. MD

- Use Model-Based Methods, as TUS-CPS has some counties with **small** (rep.< 100,000 pop.) or **zero** sample size

- *Borrowing strength* from relevant sources (Census/ Admin. info, related surveys)

- Borrowed strength from covariates & other counties with **similar** features

- The pool of auxiliary variables include:
  - 30 *county-level* demographic & socio-economic variables from ACS 2005-12, Census 2000 & 2010, & other admin. records
  - 5 *state level* tobacco policy data (cigarette taxes, clean air laws, tobacco control funding, Medicaid coverage for tobacco-related treatment, year that Quitline service initiated)

APPLICATIONS: Current Smoking Map
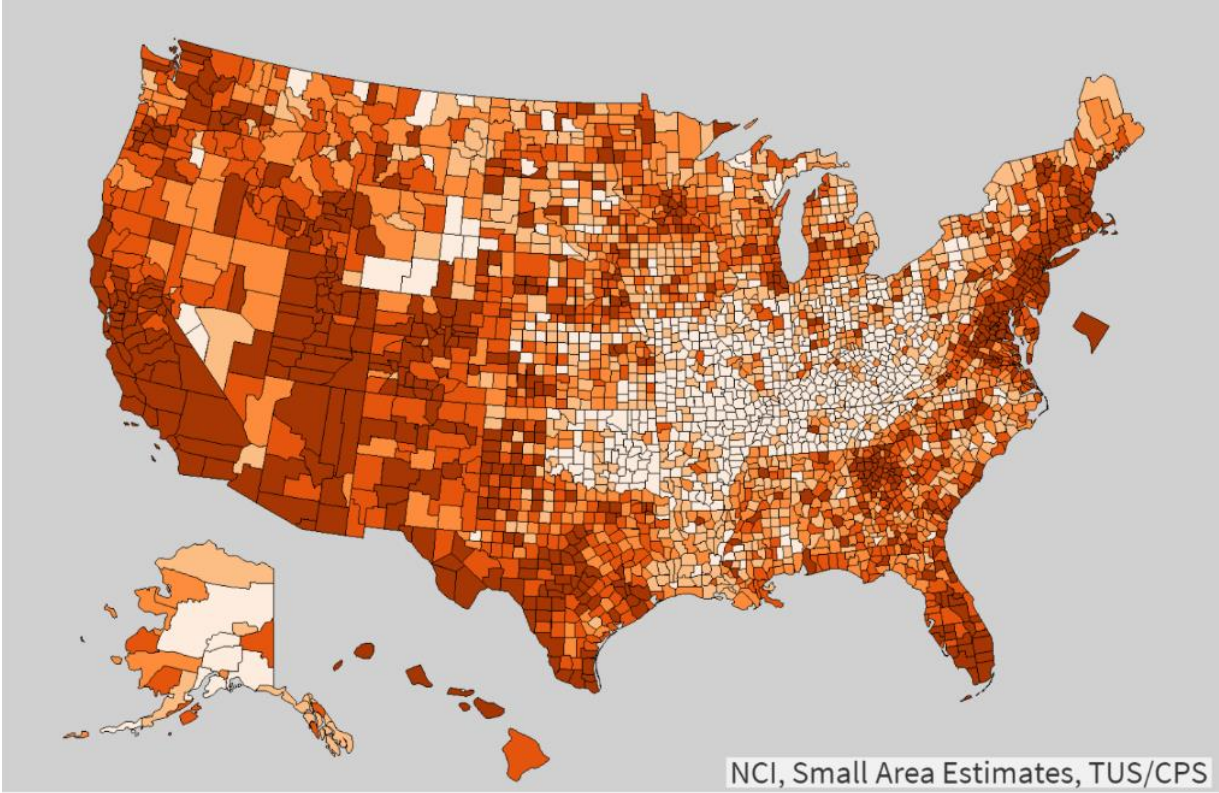
Current Smoking Prevalence (Age 18+)
US by County, 2010-2011

NCI, Small Area Estimates, TUS/CPS

Legend: 1.29 - 16.15 | > 16.15 - 18.72 | > 18.72 - 20.91 | > 20.91 - 23.8 | > 23.8 - 38.51
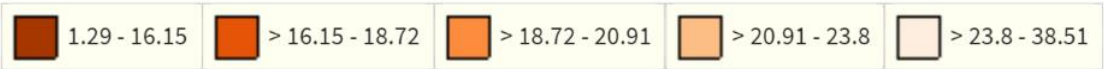
APPLICATIONS: Home Rule Map
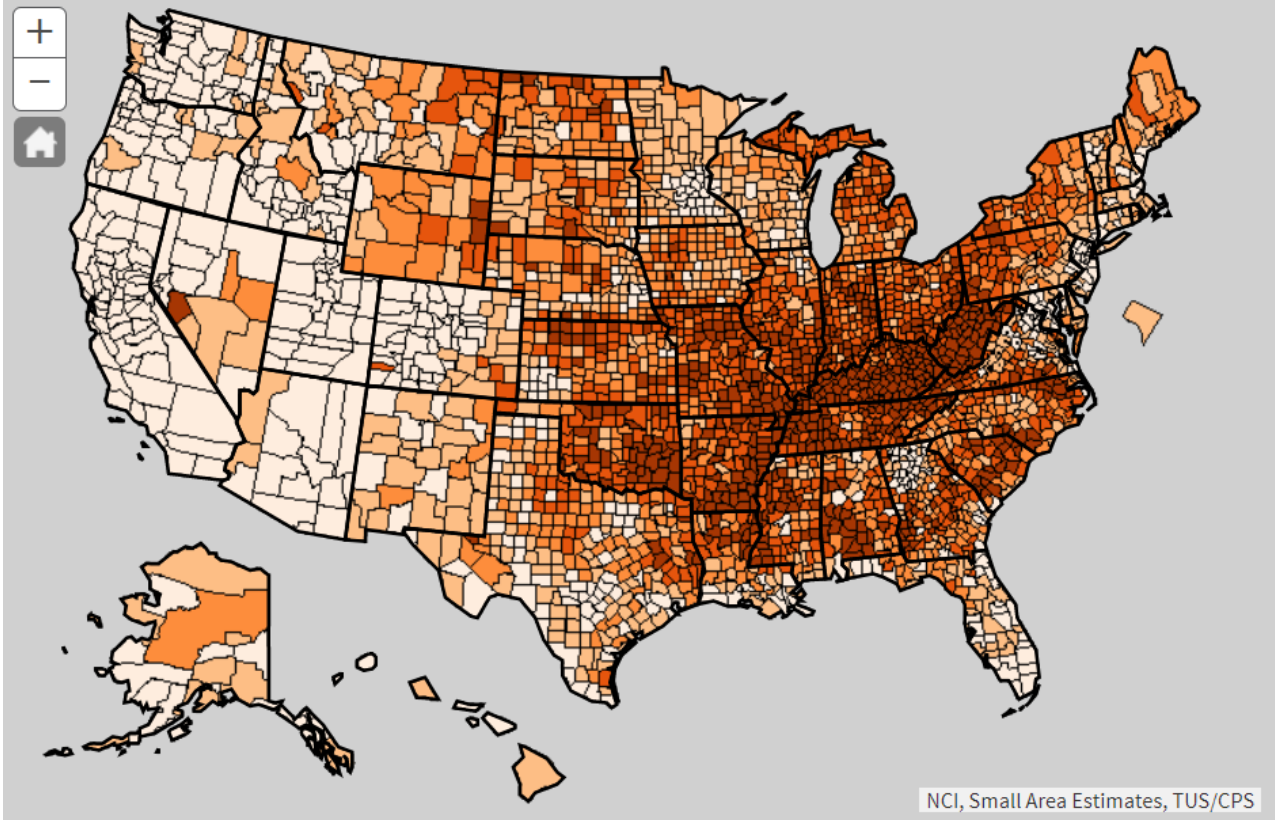
Smoke-Free Home (Age 18+)
US by County, 2010-2011

| ■ 15.35 - 65.65 | ■ > 65.65 - 70.45 | ■ > 70.45 - 74.78 | □ > 74.78 - 80.33 | □ > 80.33 - 98.69 |
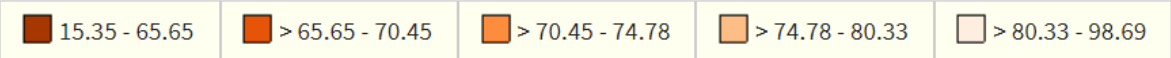
APPLICATIONS: Smoke-free Workplace Policy Maps - 1st TUS Modeled Reports & 2nd TUS State Reports from State Cancer Profiles*

*(statecancerprofiles.cancer.gov/map/map.noimage.php)
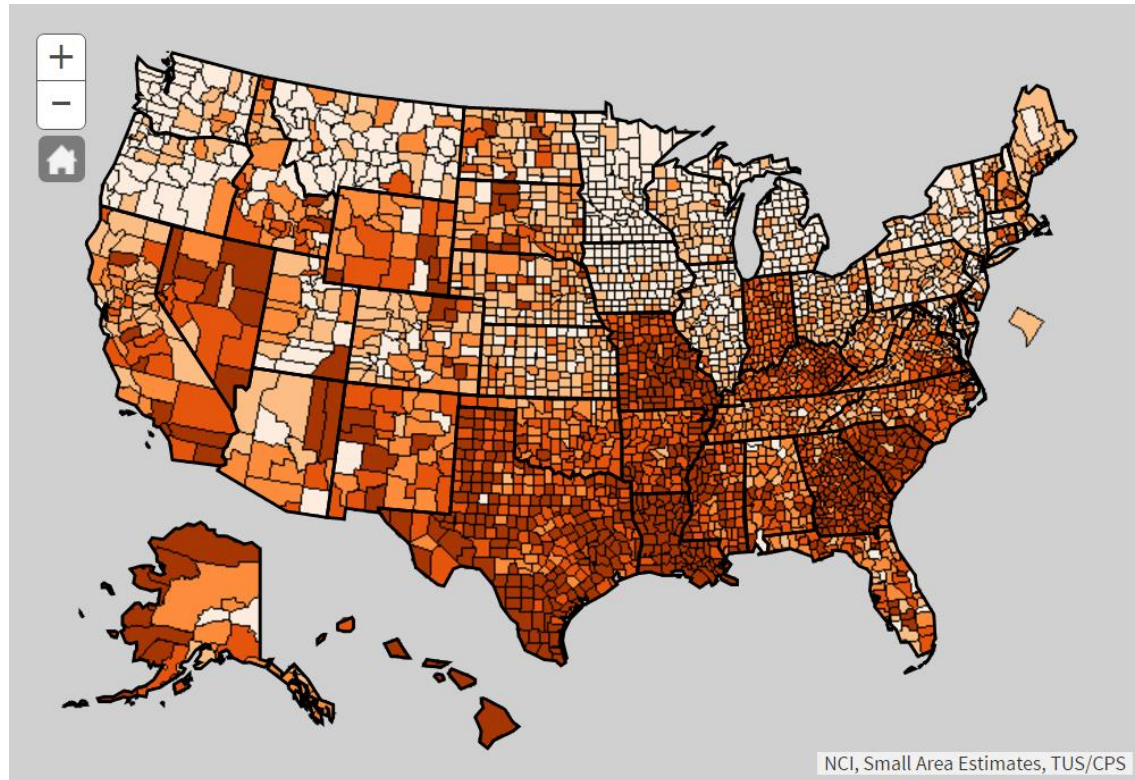
Smoke-free Workplace Policy
US by County, 2010-2011

17.38 - 76.41 | > 76.41 - 80.03 | > 80.03 - 83.79 | > 83.79 - 87.33 | > 87.33 - 100.49
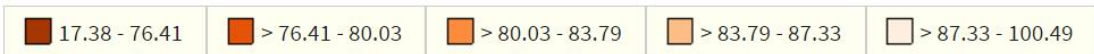
NCI, Small Area Estimates, TUS/CPS

Screening and Risk Factors for United States
(2014-2015 Tobacco Use Supplement Data)
Percent of Workers in Non-Smoking Environments (All People)
All Races (includes Hispanic), Both Sexes, Ages 18+

Percent of Workers in Non-Smoking Environments (All People) (Percent of Respondents)

Quantile Interval

62.6 to 75.9
> 75.9 to 79.5
> 79.5 to 81.8
> 81.8 to 84.9
> 84.9 to 90.0
Data Not Available

United States Rate (95% C.I.) 79.7 (79.3 - 80.1)
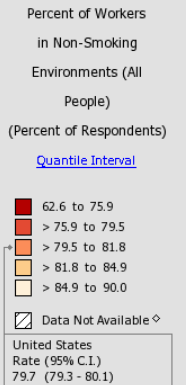
Puerto Rico

Notes:
Note: Alaska, DC, Hawaii and Puerto Rico are not drawn to scale.
Source: 2014-2015 Tobacco Use Supplement to the Current Population Survey (TUS-CPS). Detailed information about survey questions can be found here.
◊ Data not available for this combination of geography, statistic, age and race/ethnicity.
Data for the United States does not include data from Puerto Rico

MORE INFO: TUS SAE Data Accessibility & Contact

- More details & results at: **https://sae.cancer.gov/tus-cps/**

  - County level current & ever smoking prevalence estimates available upon request *

- NCI SAE general website **sae.cancer.gov**
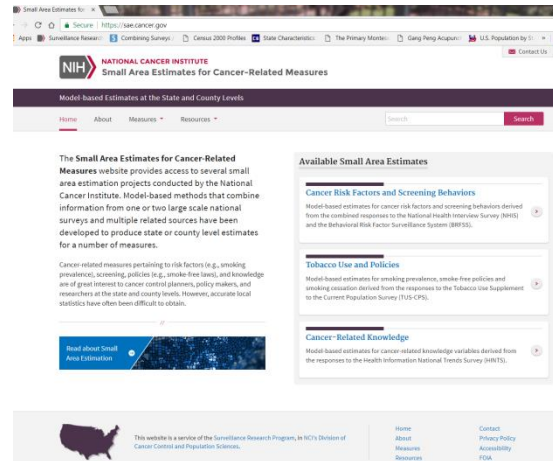
- **Future**: 2014-15 TUS wave; county estimates by race/ethnicity

- **Contact** Benmei Liu: **liub2@mail.nih.gov;** our website-SAE details PPT

* avoid confusion as similar estimates for these 2 items derived from combining NHIS & BRFSS released at https://sae.cancer.gov/nhis-brfss/

TUS-CPS Unique Characteristics & Linkages Harmonization Process

Carolyn Reyes-Guzman
NCI, DCCPS

James 'Todd' Gibson
Information Management Services, Inc

February 20, 2019

NIH NATIONAL CANCER INSTITUTE

Tobacco Use Supplement
Current Population Survey

Effort Behind Data Harmonization

- Researchers who previously wanted to examine long-term data had to track different variable names over several survey waves for a specific question of interest

- Questions on special topics with smaller sample sizes were challenging to study, creating an opportunity for using harmonized data to pool samples over time

- Data harmonization involved creating one variable name to capture all survey waves, using a "flag" variable to track the survey year

- "One-stop shopping:" harmonized public use file will be posted on the TUS-CPS website along with supporting documentation (data dictionary, dataset contents, descriptive tables and technical documentation)

Overview of TUS-CPS Harmonized Data File

- Variables:

  - Current Population Survey (CPS) Core:

    - Region, state, county (limited), family income, metropolitan status, age, sex, race*, Hispanic origin, marital status, education, employment status, nativity (respondent, mother, father)

  - Tobacco Use Supplement (TUS):

    - Cigarette Use – including menthol since 2003

    - Workplace and home smoking restrictions

    - Attitudes towards smoke-free policies in indoor work areas/public places

    - Advice by physician/dentist

    - Health perceptions/beliefs (harm reduction)

    - Smoking history, cessation, former smokers

    - Other tobacco products: cigar, pipe and smokeless tobacco use

Data harmonization processes

- Criteria for inclusion: question asked in >1 survey*

- Harmonized file includes adult (18+) self-reported responses, no proxy responses, as they bring greater reliability issues

- Special variable considerations (e.g., when item wording changed over time or when the universe of respondents <u>slightly</u> varied over time)

- Harmonized variable names capture a single question over multiple years

- Variable crosswalk provides detailed notes for the data user on rationale for harmonized variables for special considerations

*few variables were excluded due to <u>substantial</u> question wording changes or difficulty harmonizing

TUS Harmonized Codebook Screenshot

Question/Item	Harmonized Variable Name	Individual Wave Variable Name								Notes
		1992-93	1995-96	1998-99	2001-02	2003	2006-07	2010-11	2014-15	
Core CPS Items										
Survey wave	SURWAVE	1	2	3	4	5	6	7	8	Variable was constructed.
Month of interview	SURMONTH	H_MONTH	HRMONTH	HRMONTH	HRMONTH	HRMONTH	HRMONTH	HRMONTH	HRMONTH	
Year of interview	SURYEAR	H_YEAR	HRYEAR	HRYEAR4	HRYEAR4	HRYEAR4	HRYEAR4	HRYEAR4	HRYEAR4	
Family income	FAMINC	H_FAMINC	HUFAMINC	HUFAMINC	HUFAMINC	HUFAMINC	HUFAMINC	HUFAMINC	HUFAMINC	Slight changes in response categories over time. Family income allocation flag (HXFAMINC) in 2010-2011 and 2014-2015.
Number of persons in household	HHNUM	H_NUMPER	HRNUMHOU	HRNUMHOU	HRNUMHOU	HRNUMHOU	HRNUMHOU	HRNUMHOU	HRNUMHOU	
Region	REGION	HG_REG	GEREG	GEREG	GEREG	GEREG	GEREG	GEREG	GEREG	
State	FIPS	HG_FIPS	GESTFIPS	GESTFIPS	GESTFIPS	GESTFIPS	GESTFIPS	GESTFIPS	GESTFIPS	
State name	STATE		GECO	GECO	GECO	GECO	GTCO	GTCO	GTCO	Variable was constructed. This code must be used with a state code (FIPS) in order to uniquely identify a county. Not all counties identified. Constructed variable in harmonized data file is STCOUNTY. This variable includes state and county.
Metropolitan status	METSTAT	H_METSTA	GEMETSTA	GEMETSTA	GEMETSTA	GEMETSTA	GEMETSTA	GEMETSTA	GEMETSTA	
Age	AGE	A_AGE	PEAGE	PEAGE	PEAGE	PRTAGE	PRTAGE	PRTAGE	PRTAGE	Slight changes in response categories over time.
Sex	SEX	A_SEX	PESEX	PESEX	PESEX	PESEX	PESEX	PESEX	PESEX	
Race (1992-2002)	RACE92	A_RACE	PERACE	PERACE	PERACE					Only 4-5 response categories were available to respondents.
Race (2003-2015)	RACE03					PTDTRACE	PTDTRACE	PTDTRACE	PTDTRACE	21+ response categories were available to respondents.
Hispanic or non-Hispanic origin	HISPAN	A_REORGN	PRHSPNON	PRHSPNON	PRHSPNON	PEHSPNON	PEHSPNON	PEHSPNON	PEHSPNON	1992-1993 asked specificity of origin; 1995+ asked as dichotomous variable
Marital Status	MARITAL	A_MARITL	PRMARSTA	PRMARSTA	PRMARSTA	PRMARSTA	PRMARSTA	PRMARSTA	PRMARSTA	
Highest Level of School Completed or Degree Received	EDU	A_HGA	PEEDUCA	PEEDUCA	PEEDUCA	PEEDUCA	PEEDUCA	PEEDUCA	PEEDUCA	question was not asked in 2003 but variable was constructed.
Monthly Labor Force Recode	EMPLSTAT	A_LFSR	PEMLR	PEMLR	PEMLR	PEMLR	PEMLR	PEMLR	PEMLR	1992-1993 had slight differences in response categories compared with other waves.
Country of Birth	NATIVITY		PENATVTY	PENATVTY	PENATVTY	PENATVTY	PENATVTY	PENATVTY	PENATVTY	
Mother's Country of Birth	NATIVITYM		PEMNTVTY	PEMNTVTY	PEMNTVTY	PEMNTVTY	PEMNTVTY	PEMNTVTY	PEMNTVTY	
Father's Country of Birth	NATIVITYF		PEFNTVTY	PEFNTVTY	PEFNTVTY	PEFNTVTY	PEFNTVTY	PEFNTVTY	PEFNTVTY	
Citizenship Status	CITIZEN		PRCITSHP	PRCITSHP	PRCITSHP	PRCITSHP	PRCITSHP	PRCITSHP	PRCITSHP	Slight changes in response categories over time.

Examples of Decision-Making Scenarios

- Former smokers section: the universe of respondents slightly changed over time; some years asked "in the last 3 years", whereas later years asked "in the last 5 years."

  - To harmonize, the length of quit (LOQ) was used to set the universe to include only the last 3 years

- Length of quit items: LOQ in earlier surveys was truncated at 12 months, but not truncated in later years.

  - To harmonize, later survey years were truncated to 12 months

Examples of Decision-Making Scenarios

- Quit attempts section: earlier years first asked "Have you ever made a serious attempt to stop smoking…" then "During the past 12 months, have you made a serious attempt to stop smoking…", but later years asked in the reverse.

    - It isn't clear whether the order could have impacted patterns of responses.

    - To harmonize, the reverse order was imputed and a note added to the data dictionary to clarify.

Examples of Decision-Making Scenarios

- Ever and current use of any other tobacco products (e-cigarette items were asked starting in 2014)

  - Without more years of data, the variable is not as useful so these items were thus dropped

- Time to first use of a non-cigarette tobacco product: a summary variable was deemed too difficult to harmonize.

  - This will be a follow-up item once the 2018-19 data are available for possible harmonization then.

Overview of Tobacco Use Supplement (TUS) to the Current Population Survey (CPS) Harmonized Data File

Overview of TUS-CPS Harmonized Data File

- Data from the 9 survey waves:

  1. 1992-1993: September 1992, January 1993 and May 1993

  2. 1995-1996: September 1995, January 1996 and May 1996

  3. 1998-1999: September 1998, January 1999 and May 1999

  4. 2000: January 2000 and May 2000

     - January and May 2000 survey was done because of change in wording in 1998-99. Limited number of questions asked. Additional information available in the 2000 technical use documentation.

  5. 2001-2002: June 2001, November 2001 and February 2002

  6. 2003: February 2003, June 2003 and November 2003

  7. 2006-2007: May 2006, August 2006 and January 2007

  8. 2010-2011: May 2010, August 2010 and January 2011

  9. 2014-2015: July 2014, January 2015 and May 2015

Overview of TUS-CPS Harmonized Data File

- **January and May 2000**

  - The September 1998, January 2000 and May 2000 supplements have different wording for the "other tobacco use" questions than the 1992-93, 1995-96 and January and May 1999

  - In September 1998, the questions were changed to more appropriately capture some day/occasional usage of other tobacco products by changing the previous wording

    From: (Have/Has) (you/name) EVER used pipes, cigars, chewing tobacco or snuff *on a regular basis*?"
    To: (Have/Has) (you/name) EVER used pipes, cigars, chewing tobacco or snuff?"

  - In January and May 1999 the wording reverted back to "(Have/Has) (you/name) EVER used pipes, cigars, chewing tobacco or snuff *on a regular basis*?"

  - It is recommended that for more accurate estimates of other tobacco products the January 2000 and May 2000 data should be averaged or January 2000 and May 2000 be averaged with September 1998 other tobacco product data.

Overview of TUS-CPS Harmonized Data File

From TUS-CPS Website:

Race/Ethnicity changes in 2003 CPS: In 2003, significant changes were made to race/ethnicity questions in the CPS. Respondents were able to select more than one race when answering the survey. This change in wording does not impact smoking estimates and trends made for the entire nation from the TUSCS-CPS, but it could potentially impact smoking estimates and trends made by race/ethnicity. NCI has developed a method to construct single race-estimates using data from the post-2003 TUSCS-CPS. The method is useful when trends over time are being examined for single race groups using both pre-2003 and post-2003 data. More information is available in the Bridging Estimates by Race report, (PDF) which describes the method and gives an initial assessment of the usefulness of the race adjustment. Also, see Trends in Smoking Prevalence by Race based on the Tobacco Use Supplement to the Current Population Survey (PDF) for an application of this race bridging.

https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/race_bridging_071307.pdf
https://pdfs.semanticscholar.org/f538/2ae7f5f41a9708d12fce3530f87360088b6b.pdf

Overview of TUS-CPS Harmonized Data File

RACE92: 1992-2002	RACE03: 2003-2015
1 = White	1 = White only
2 = Black	2 = Black only
3 = American Indian, Aleut, Eskimo	3 = American Indian, Alaskan Native only
4 = Asian or Pacific Islander	4 = Asian only
5 = Other (1992-93 only, 0.3% of sample)	5 = Hawaiian/Pacific Islander only
	6 = White and Black
	7 = White and American Indian, AK Native
	.
	.
	19 = W-B-AI-A
	20 = Other 2 or 3 races
	21 = Other 4 races or all 5 races

Overview of TUS-CPS Harmonized Data File

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: RACE03

	2003		2006-2007		2010-2011		2014-2015	
	Sample	Percent	Sample	Percent	Sample	Percent	Sample	Percent
Race (2003 and later)								
White only	157,192	85.52	146,230	85.01	143,069	83.49	135,695	82.78
Black only	16,148	8.79	15,178	8.82	16,778	9.79	16,929	10.33
American Indian, Alaskan Native only	1,877	1.02	1,710	0.99	1,705	0.99	1,870	1.14
Asian only	5,379	2.93	5,583	3.25	6,613	3.86	6,549	4.00
Hawaiian/Pacific Islander only	448	0.24	496	0.29	617	0.36	570	0.35
White and American Indian, Alaskan Native	1,700	0.92	1,591	0.92	1,232	0.72	997	0.61
Other 2 or more races	1,066	0.58	1,235	0.72	1,351	0.79	1,310	0.80
Total	183,810	100.00	172,023	100.00	171,365	100.00	163,920	100.00

Overview of TUS-CPS Harmonized Data File

- Self respondents, Ages 18 and over

- Weights:

  - Main self-response weight included

    - Controlled for age, race, sex, and Hispanic origin and individual state.

    - Weights can be used to generated estimates for each or the 26 surveys (Sept 1992, Jan 1993,…,July 2015) individually, or for

    - Most reliable estimates, especially for states or sub-state level, will come from combining the data from the 2 surveys in 2000 and all 3 collection periods for all other survey waves. When combining data the weight needs to be divided by number of surveys being combined.

  - Replicate weight will be released in the future as a separate file. Code and information on linking with main harmonized data file and use will be supplied with the file.

Overview of TUS-CPS Harmonized Data File

- **File specifics:**
  - Filename: harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.dat
  - ASCII Text file
  - Number of records: 1,593,413
  - Record Length: 554
  - Number of variables: 225 (199 TUS)

Heterogeneity in tobacco use behaviors among US black populations

Kelvin Choi, PhD, MPH

Launick Saint-Fort, BS

Division of Intramural Research, NIMHD

Understanding Tobacco Use Trends: Leveraging Harmonized Data from the U.S.
Tobacco Use Supplement – Current Population Survey, 1992-2015
Wednesday, February 20, 2019
8:30 am – 11:30 am PST

Tobacco Use Supplement
Current Population Survey

National Institute
on Minority Health
and Health Disparities

Disclaimer

The opinions and comments expressed during this presentation are the presenter's own and do not necessarily represent those of the National Institute on Minority Health and Health Disparities, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, and the U.S. Government.

Background

- Smoking remains the primary cause of death of African-Americans in the U.S.

- 3.8 million foreign-born blacks are currently living in the U.S.
  - 8.7% of the U.S. black population in 2013 and 16.5% by 2060

- Previous studies found differences in prevalence of smoking by global region of origin and age of immigration

Objectives

- Explore whether there is heterogeneity in tobacco-use behaviors (including current cigarette and cigar smoking, time to first cigarette, and menthol cigarette use) among U.S. blacks by global region of origin and age at immigration.

Tobacco Use Supplement
Current Population Survey

Problem!

US population = 100 people

Blacks = 20 people

Foreign-born blacks < 2 people!

Country of birth!

Available in NHIS restricted-use file

TUS-CPS to the rescue!

- Large sample sizes
  - >100,000 per data collection wave

- Country of birth
  - Public use file

- Year of entry to the US
  - Calculate age of immigration

Methods

- Data: 2006-07, 2010-11, and 2014-15 Current Populations Survey - Tobacco Use Supplement (CPS-TUS)

- Sample (n=47,857):
  - Self-identified blacks (regardless of Latino origin)
  - 18 years or older
  - Self-respondents
  - Global region of birth
    - U.S. (including Guam, Puerto Rico, the U.S. Virgin Islands, and other U.S. island areas): N=43,560
    - Africa (including both North and Sub-Saharan Africa): N=1,911
    - West Indies: N=2,194
    - Europe: N=192

Country of birth

ATTACHMENT 14

COUNTRIES AND AREAS OF THE WORLD

Current Population Survey

Starting May 2012

Code	Name	Code	Name
057	United States	158	Armenia
060	American Samoa	159	Azerbaijan
066	Guam	160	Belarus
069	Northern Marianas	161	Georgia
073	Puerto Rico	162	Moldova
078	U. S. Virgin Islands	163	Russia
100	Albania	164	Ukraine
102	Austria	165	USSR
103	Belgium	166	Europe, not specified
104	Bulgaria	168	Montenegro
105	Czechoslovakia	200	Afghanistan
106	Denmark	202	Bangladesh
108	Finland	203	Bhutan
109	France	205	Myanmar (Burma)
110	Germany	206	Cambodia

Methods

- Measures:
  - Age at entry = Year of survey – Age – Year of entry
  - Current tobacco use = some days + every day
    - Cigarettes
    - Cigars
  - Established smokers (100 cigarettes in the lifetime)
    - Former smoker
    - Started regular smoking before 18 years old: Age of regular smoking
  - Current smokers (Some days + every day cigarette smoking)
    - First cigarette within 30 minutes after waking
    - Regular menthol cigarette smoking

Methods

- Measures (continued)
  - Demographics: Survey year, age, gender, educational attainment, income, and US census region.
- Analysis
  - Weighted multivariable logistic regression

Results: Full sample
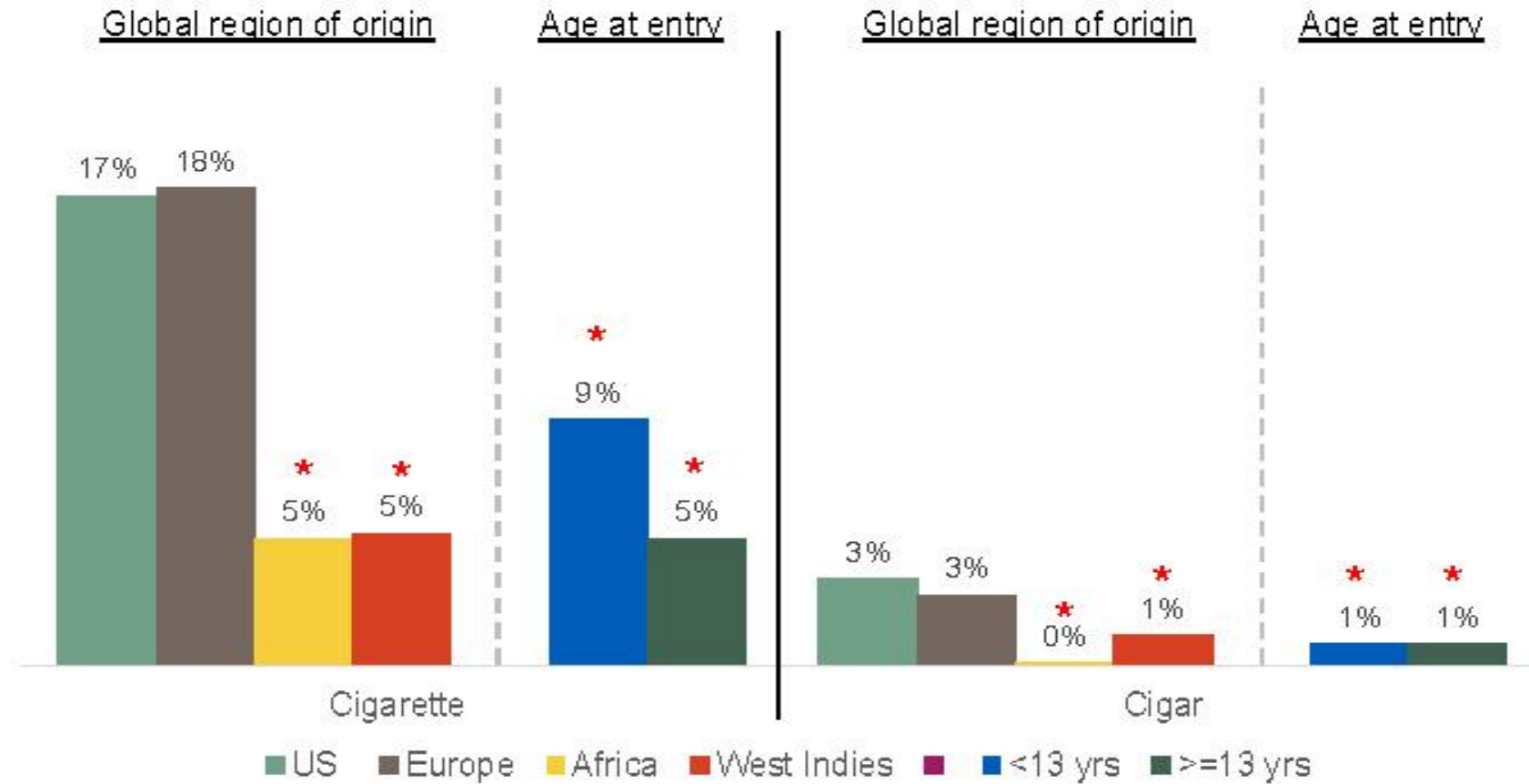
(Reference: US-born; *=p<0.05)

Results: Established smokers
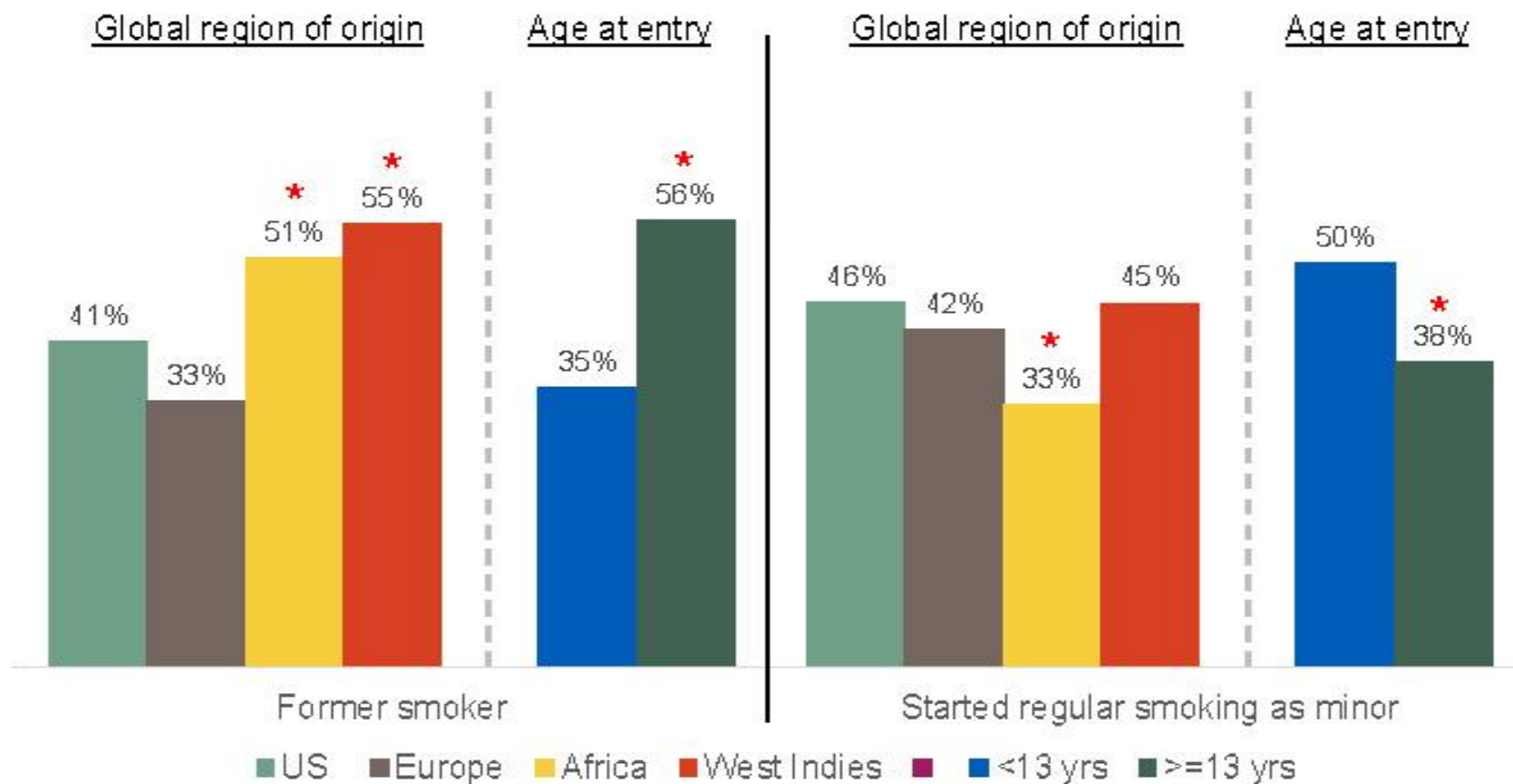
(Reference: US-born; *=p<0.05)

Results: Current smokers
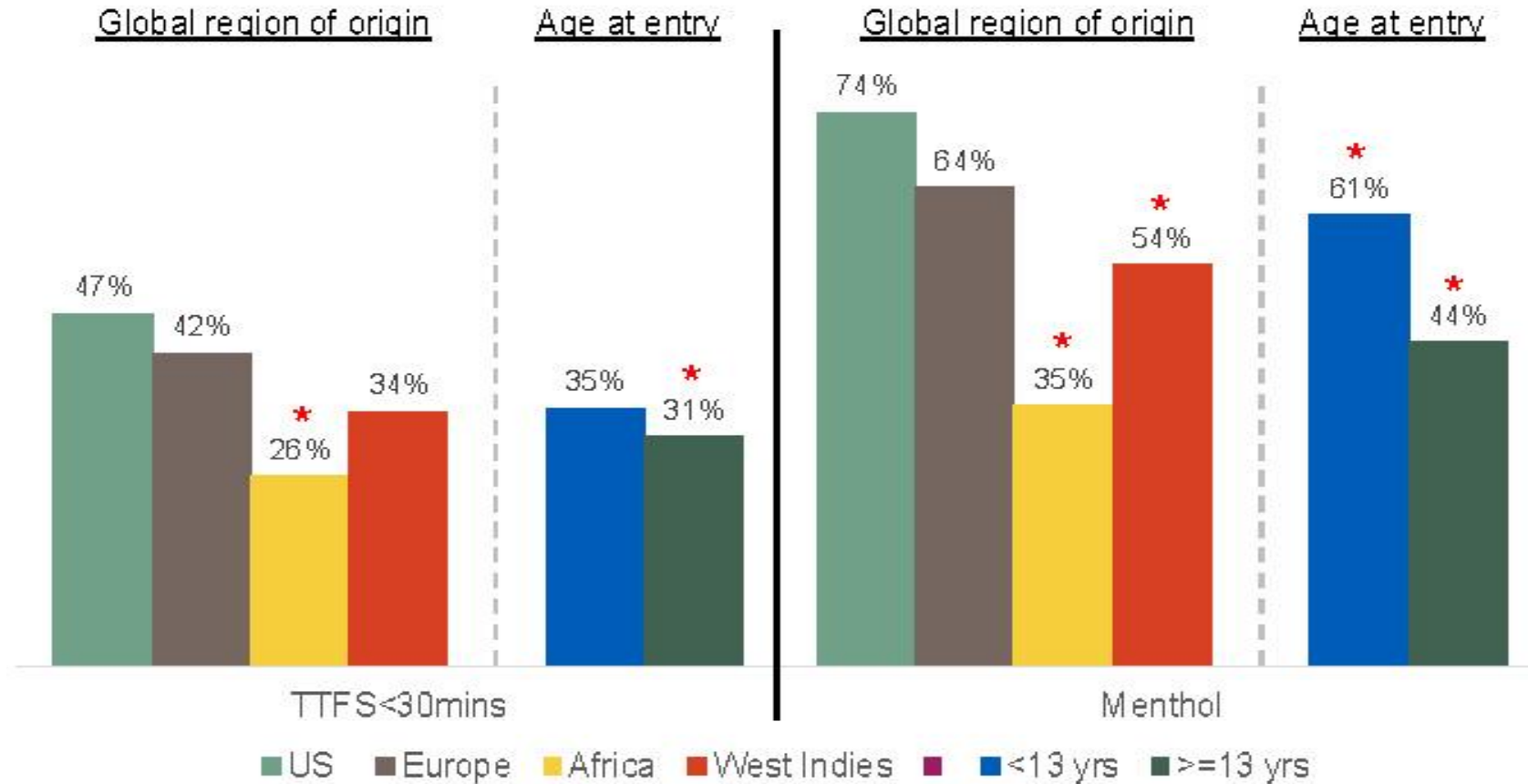
(Reference: US-born; *=p<0.05)

Conclusions

- TUS-CPS harmonized data provide great opportunities to study racial/ethnic disparities in tobacco use!

- Heterogeneity by global region of origin
  - Not all blacks are the same
  - Menthol cigarette smoking probably is a result of targeted marketing, not preference

Questions?

Kelvin.choi@nih.gov

Saint-Fort, L., Choi, K. Heterogeneity in Tobacco-Use Behaviors among U.S. Blacks by Global Region of Origin. Journal of Immigrant and Minority Health. In-press.

Mortality Analysis Using Harmonized Tobacco Use Data

Maki Inoue-Choi, PhD, MS

Neal D Freedman, PhD, MPH

February 20, 2019

NIH NATIONAL CANCER INSTITUTE

Tobacco Use Supplement
Current Population Survey

Falling prevalence of US cigarette smoking
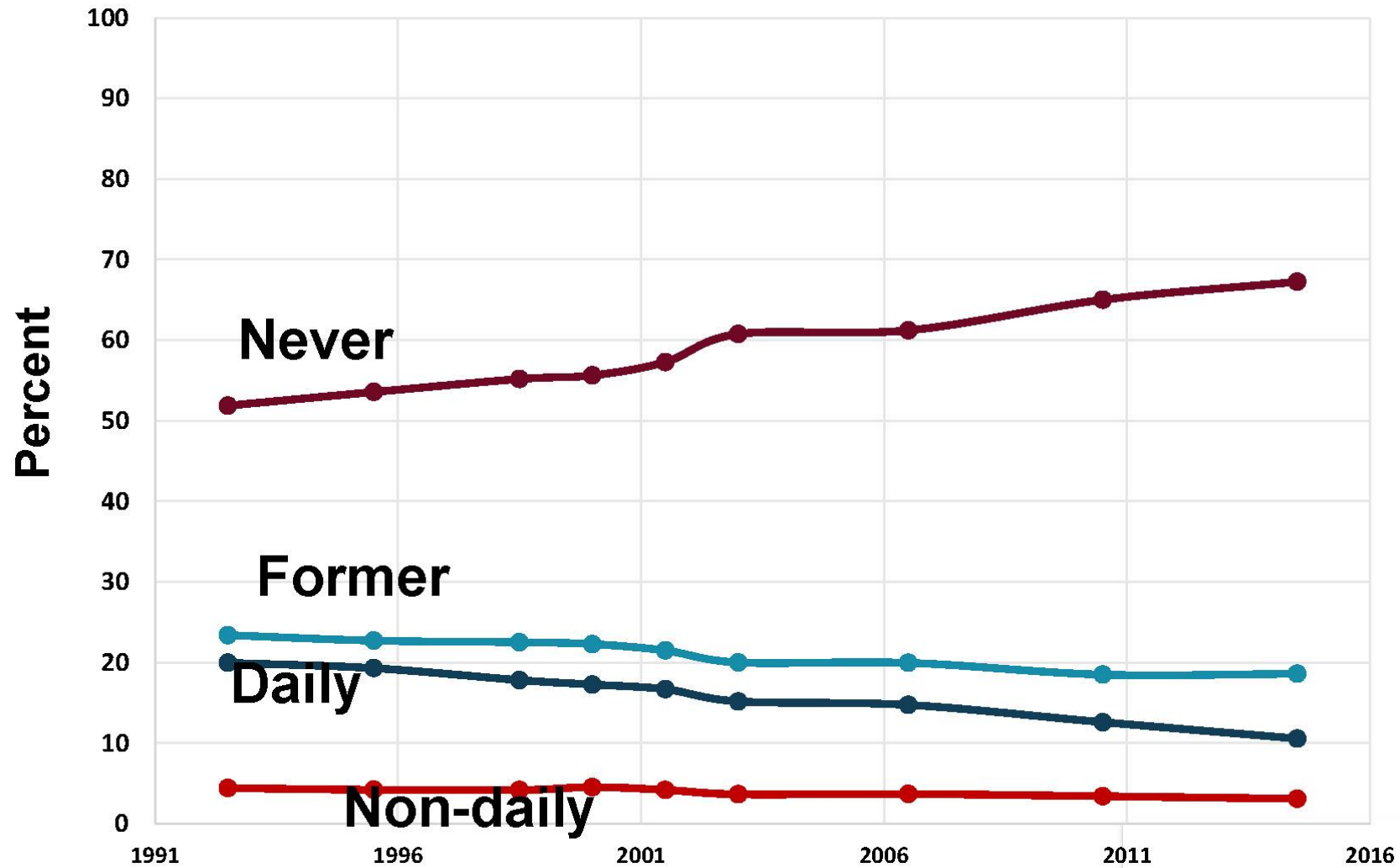
TUS-CPS NLMS mortality linkage presents unique opportunities for studies of mortality

- Detailed questions about range of tobacco products sold in the US

  detailed usage patterns, lifetime history, cessation

- Nationally representative, so reflects US population and products used

  wide age range, birth cohorts, and race/ethnicity group

NIH NATIONAL CANCER INSTITUTE

US prevalence of different tobacco products

	Any	Cigs	Cigars	Hookah /pipe	E-cig	Smoke-less	≥2 types
Adults	20.1%	15.1%	3.4%	1.2%	3.5%	2.3%	3.9%
High school	20.2%	8.0%	7.7%	6.2%	11.3%	5.8%	9.6%

- **Non-daily smokers**

NHIS, *MMWR*, 2017

Linked Mortality Data via National Longitudinal Mortality Study (NLMS)

- NLMS
  - Socioeconomic data collected as a part of CPS by US Census Bureau
  - Cause of death information from National Death Index through 2011
- Linked to a subset of TUS-CPS
- Linked mortality data currently available via RDC or Census

Mortality Analysis Example 1

Cigarette, Cigar, and Pipe Use and All-Cause and Cause-Specific Mortality, 1985-2011

Understanding the Health Risks of Non-cigarette Tobacco Product Use

- Because trends are recent, most epidemiologic studies lack users and detailed assessment

- Harmonized tobacco use data in the 1992-1993, 1995-1996, 1998-1999, 2000, 2001-2002, 2003, 2006-2007, and 2010-2011 TUS-CPS and the 1985 CPS

Methods (1)

- 357,420 NLMS participants
- Exclusive use of cigarettes, cigars, and pipes
  - Ever users: Smoked 100+ cigarettes in the lifetime
  - Former/current; daily/non-daily (some day)
- All-cause mortality (n = 51,150)
- Cause-specific mortality

  tobacco-related cancer; circulatory disease; cardiovascular disease; cerebrovascular disease; respiratory disease; diabetes

Methods (2)

- Cox proportional hazards proportional regression
  - Age as the underlying time metric
  - Covariates: sex, race/ethnicity, education, CPS survey year
  - Reference: never uses of cigarettes, cigars, pipes, or smokeless tobacco
- NLMS survey weights set to the non-institutional US population

Association of Exclusive Cigarette, Cigar, and Pipe Use with Tobacco-related Cancer Mortality, 1985-2011

	Exclusive cigarette		Exclusive cigar		Exclusive pipe	
	Non-daily	Daily	Non-daily	Daily	Non-daily	Daily
Total, n	9,414	57,251	608	531	78	1,099
Death, n	236	2,984	*	*	0	22
Hazard Ratio	2.31 (2.01-2.65)	4.33 (4.09-4.58)	1.08 (0.45-2.61)	1.80 (1.20-2.69)	-	1.75 (1.16 2.64)

Mortality Analysis Example 2

All-cause and Cause-specific Mortality among Non-daily Cigarette Smokers, 1992-2011

National Health Interview Survey (NHIS)

National Health Interview Survey

The principal source of information on the health of the U.S. population

NATIONAL HEALTH INTERVIEW SURVEY

Since 1957

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

CDC
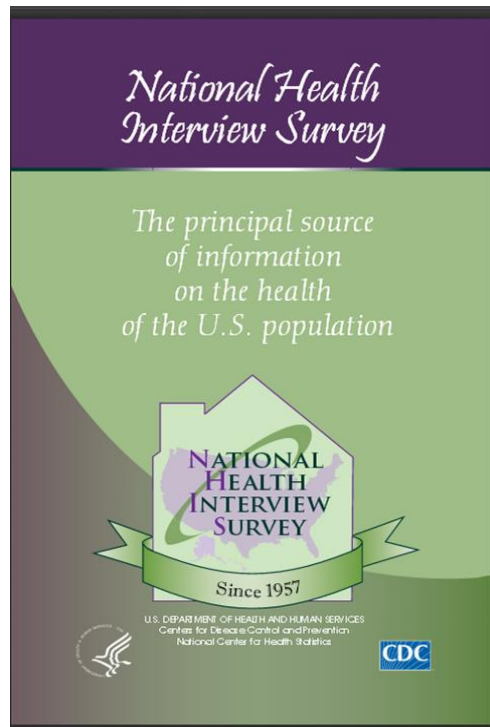
- Annual household survey

- Nationally representative

- Tobacco data collected in Supplements since 1965

- Linkage to mortality data from National Death Index (NDI)

National Health Interview Survey

Lifelong non-daily smokers had higher overall mortality risk: NHIS 1991, 1992 & 1995

Total N = 70,913

Daily smokers	30 days	600 cigarettes
Non-daily smokers	15 days	50 cigarettes

	Never	**Daily**	**Lifelong Non-daily**
Hazard Ratio	1.00 (reference)	2.50 (2.35-2.66)	1.72 (1.36-2.18)

Non-daily smokers lose 5 years on average of lifetime vs. never smokers: The NHIS 1991, 1992, and 1995

NATIONAL CANCER INSTITUTE
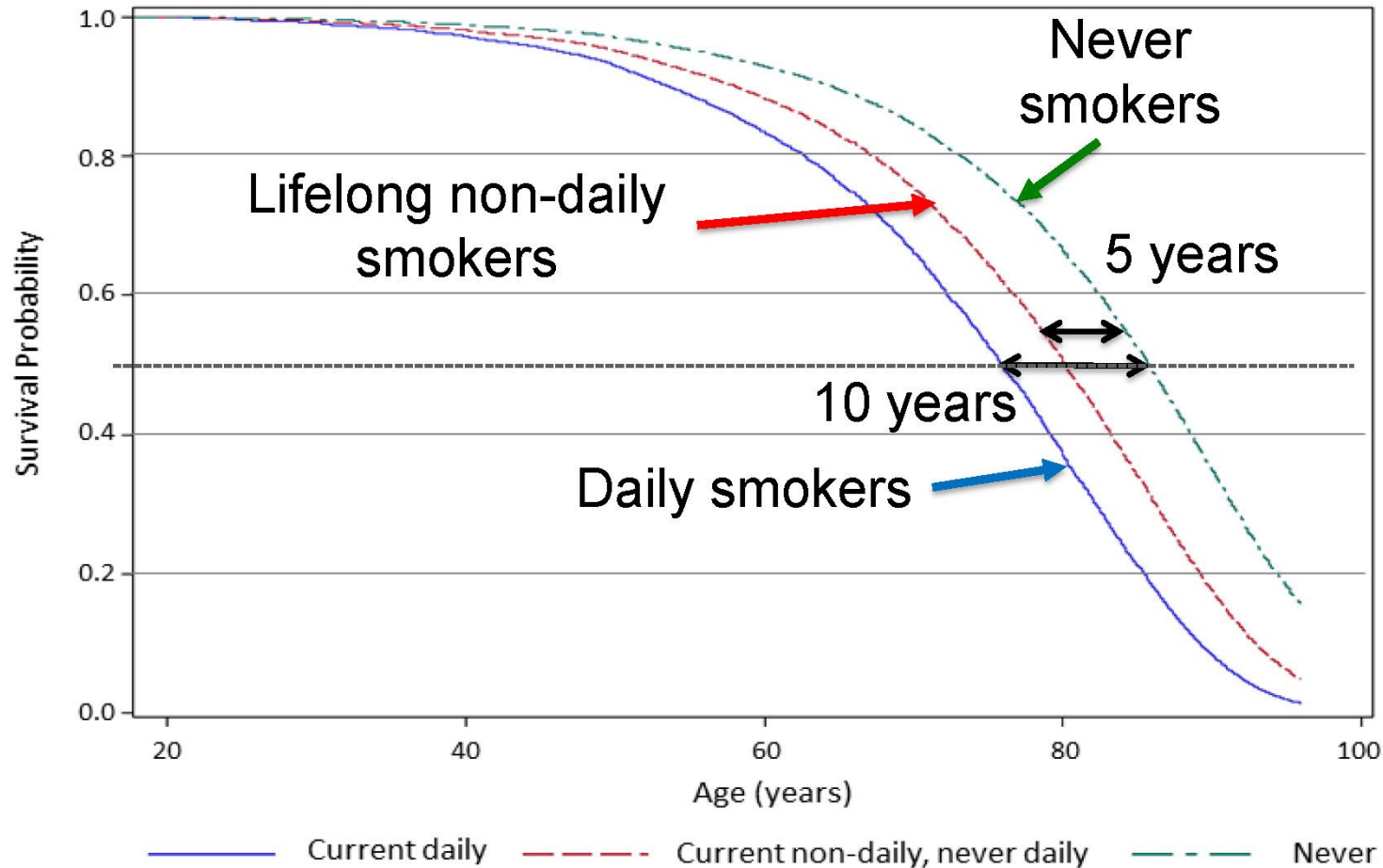
Limitations of previous analyses

- Lack of statistical power detailed information on usage patterns

  - Number of cigarettes smoked per month

  - Reducing from daily to non-daily smoking

- Cause-specific mortality

Methods

- ~505,000 NLMS participants

- Harmonized tobacco use data in the 1992-1993, 1995-1996, 1998-1999, 2000, 2001-2002, 2003, 2006-2007, and 2010-2011 TUS-CPS

- Cigarette use
  - Ever users: Smoked 100+ cigarettes in the lifetime
  - Former/current (past 30 days); daily/non-daily (some day)
  - Number of cigarettes smoked per day/on days they smoked
  - Number of days smoked in the past 30 days
  - Age started smoking; age at cessation; years since reducing from daily to non-daily smoking

All-cause Mortality by Daily and Non-daily Tobacco Use, 1992-2011

Manuscript in preparation; **Poster #43, Session 4**   82
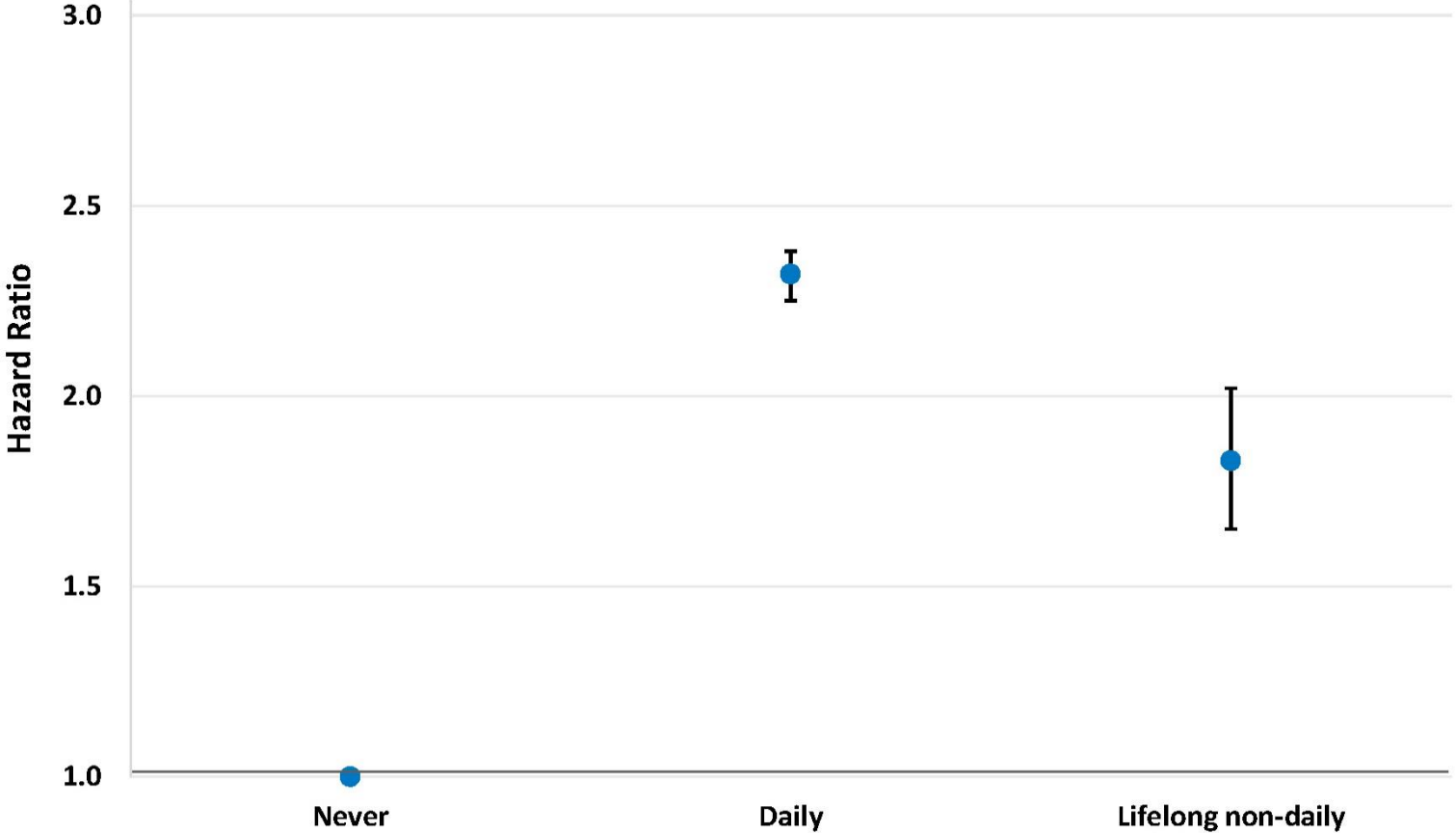
All-cause Mortality by Years since Reducing from Daily to Non-daily Smoking, 1992-2011
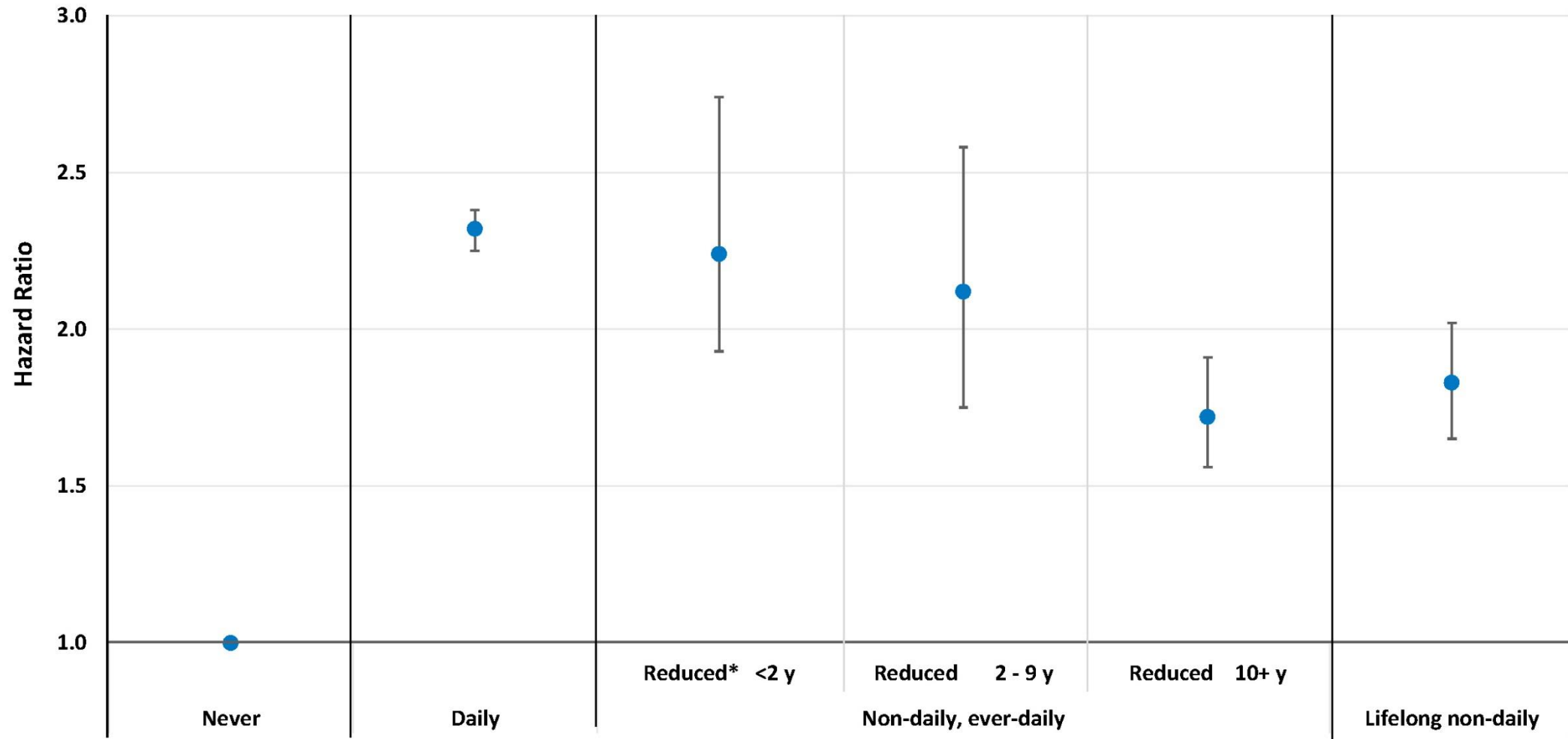
All-cause Mortality by Number of Cigarettes Smoked per Month among Daily Smokers, 1992-2011

Manuscript in preparation; **Poster #43, Session 4**   84
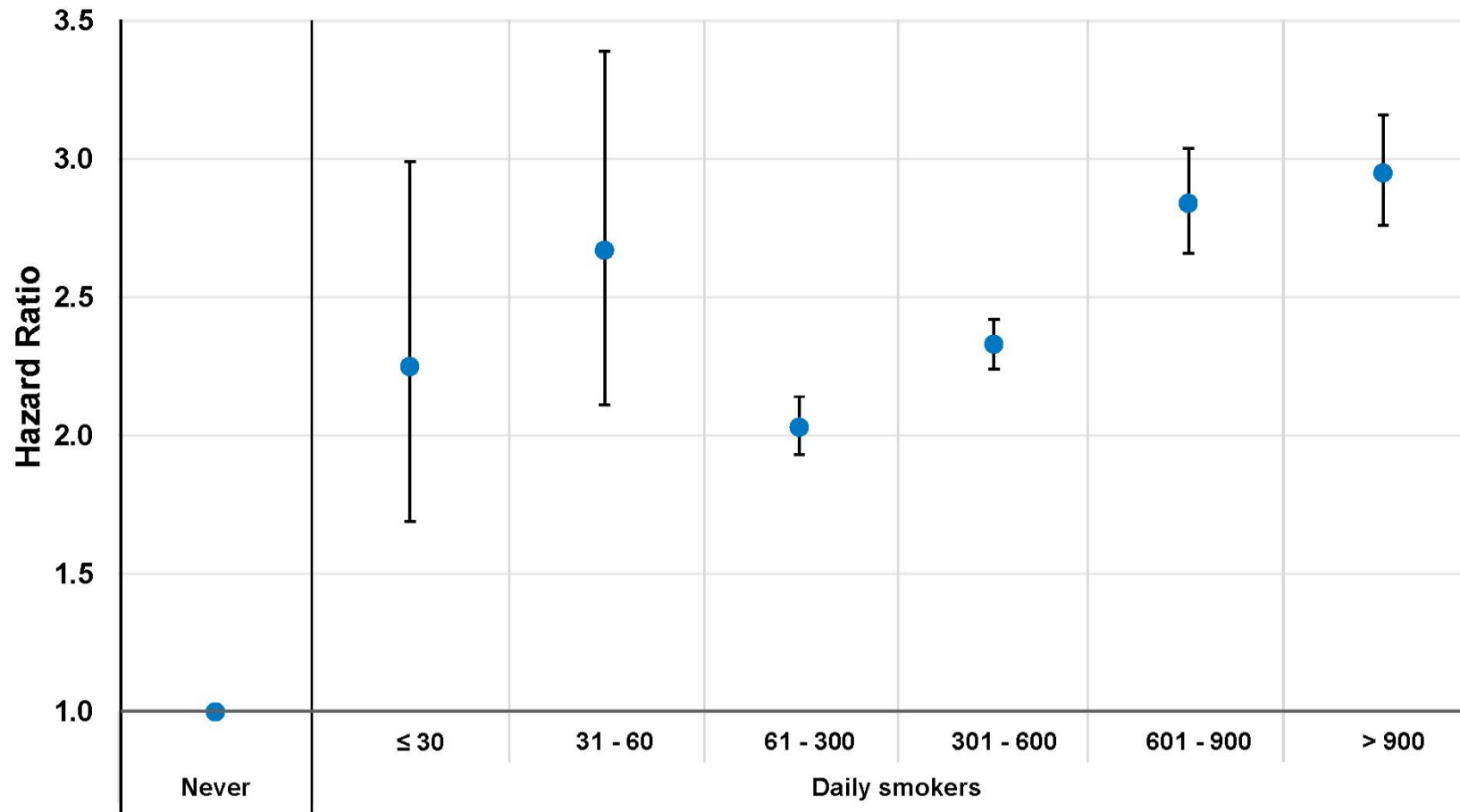
All-cause Mortality by Number of Cigarettes Smoked per Month among Non-daily and Daily Smokers, 1992-2011
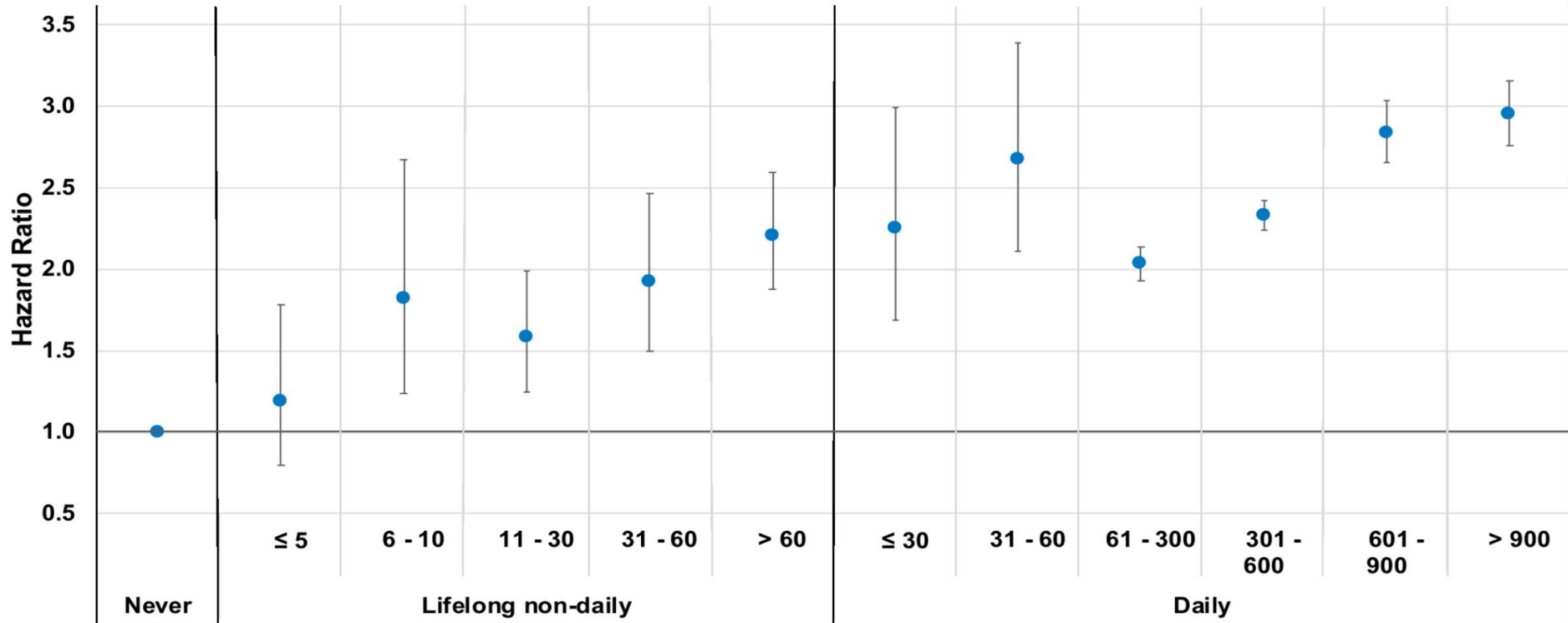
Summary

- Detailed tobacco usage data available in the harmonized TUS-CPS data set and linked mortality data are wonderful resources for mortality analyses.

- A large sample size of the harmonized TUS-CPS data enables examination of mortality risks in relation to detailed tobacco usage pattern.

- Risk estimates from the harmonized TUS-CPS data are representative of the U.S. civilian non-institutionalized adult population.

Acknowledgements

NCI/Division of Cancer Epidemiology and Genetics

    Barry I. Graubard

    Patricia Hartge

    Neil E. Caporaso

NCI/Division of Cancer Control and Population Sciences

    Carolyn Reyes-Guzman

    Anne Hartman

    Michele Bloch

FDA/Center for Tobacco Products

    Carol H. Christensen

    Brian Rostron

    Benjamin Apelberg

U.S. Census Bureau

    Candace M. Cosgrove

    Norman J. Johnson

IMS

    Timothy S. McNeel

    Todd Gibson

How to Use the Harmonized Data

James 'Todd' Gibson

Information Management Services, Inc.

February 20, 2019

NIH NATIONAL CANCER INSTITUTE

Tobacco Use Supplement
Current Population Survey

How to Use the Harmonized Data File

- Obtaining the Data

- Harmonized Data File Examples

- Hands-on Exercises

- Questions

- Contact Information

Obtaining the Data

NIH NATIONAL CANCER INSTITUTE

Obtaining the Data

- Data available on the TUS-CPS website: https://cancercontrol.cancer.gov/brp/tcrb/tus-cps

TUS–CPS Questionnaires and Data Files

Home / Program Branches / TCRB / TUS-CPS / TUS-CPS Questionnaires and Data Files

SECTION MENU

- TCRB Key Initiatives
- TCRB Staff
- TCRB Funding Opportunities
- Tobacco Use Supplement
  - TUS-CPS Survey Topics
  - TUS-CPS Questionnaires and Data Files
  - TUS-CPS Linkages
  - TUS-CPS Publications
- Tobacco Control Monograph Series

Tobacco Use Supplement
Current Population Survey

Questionnaires

The questionnaires for the 2001-2002, 2003, 2006-2007, 2010-11, and 2014-15 Tobacco Use Supplement to the Current Population Survey (TUS-CPS) are available for download below in two formats. In addition, for the May 2010-2011 special Longitudinal Cohort TUS-CPS, please see below - "2010-11 TUS-CPS" for the May 2010 baseline questionnaire and "May 2011 Follow-up" for the May 2011 follow-up questionnaire.

Data Files

SAS programs for creating a permanent SAS dataset from the .dat and ASCII text files are available for the harmonized and single wave public use files, respectively. For harmonized data files, please refer to the technical documentation Section V. For single wave public use files, please refer to the data table below. Each program reads all variables from the core and the supplement and creates a SAS dataset containing all the variables.

Now Available: Harmonized Dataset 1992-2015

Interested in tracking trends over time? How about enlarging your analytic sample size? Data from all currently available waves (1992-2015) were recently harmonized into a single dataset. Harmonized topics include cigarette use, workplace and home smoking restrictions, attitudes toward smokefree policies, advice to quit by physician/dentist, health perceptions/beliefs, and other tobacco product use.

To download data:	For variable/value listing:	For user guide:
Harmonized Dataset (TXT)	Data Dictionary (TXT)	Technical Documentation (PDF)

For variable crosswalk and notes:	For overview of database (proc) contents:	For frequency tables:
Codebook (XLSX)	Contents list (PDF)	Descriptive tables (PDF)

Obtaining the Data

- Included with the Data File

  - SAS Programs to Create a SAS dataset from the ASCII Text File

    - Main: Reads the data and creates the SAS dataset

    - Format: Formats for all variables in the dataset. Program is called by main program.

- Other Helpful Information

  - Technical Documentation: Overview of CPS, TUS and Harmonized Data

  - Proc Contents of the Data File

  - Unweighted Frequency Tables of All Variables

  - Microsoft Excel Table Listing the TUS Harmonized Variables and the Source Variable by Survey Wave.

Harmonized Data File Examples

Harmonized Data File Examples

- Example 1:

  - Example of a variable that was easy to harmonize.

- Example 2:

  - More complex harmonized variable.

- Example 3:

  - Example of a variable constructed from multiple variables.

Harmonized Data File Examples

- Example 1: Example of a variable that was easy to harmonize

  - Harmonized Variable CIG100: Smoked 100 cigarettes in entire life

  - Supplement question: Have you smoked at least 100 cigarettes in your entire life?

  - Variable names and column positions changed over time in individual survey files. User would need to read in the 26 survey files and create a new variable. 1992-93 variable coded differently than other years.

    - 1992-93: A_S32 (col 361)

    - 1995-96: PES32 (815-816)

    - 1998-99: PES32 (857-858)

    - 2000, 2001-02: PES 32 (859-860)

    - 2003: PEA1 (879-880)

    - 2006-07, 2010-11, 2014-15: PEA1 (955-956)

Harmonized Data File Examples

- Example 1:
  - SAS program code to generate unweighted table of CIG100 by survey wave

```
Data Harmon;
   Set MyData.Harmon;
   Keep SurWave Cig100 SRWeight;
Run;

Proc Tabulate Data=Harmon Missing;
   Title1 "Tobacco Use Supplement to The Current Population Survey";
   Title2 "Harmonized Data File";
   Title3 "Variable: CIG100";
   Class SurWave Cig100;
   Table Cig100 All="Total",
         SurWave=""*N="Sample"*F=Comma8.;
Run;
```

Harmonized Data File Examples

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: CIG100

	1992-1993	1995-1996	1998-1999	2000	2001-2002	2003	2006-2007	2010-2011	2014-2015
	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample
Smoked 100 cigarettes in entire life									
-9: No response	43	.	.	.	2	1	.	1	4
-3: Refused	.	59	76	53	82	236	174	174	171
-2: Don't know	625	160	241	179	242	212	237	222	183
1: Yes	110,998	88,179	80,557	56,244	81,935	74,701	70,315	63,455	58,263
2: No	116,886	98,743	95,578	68,106	103,307	108,660	101,297	107,513	105,299
Total	228,552	187,141	176,452	124,582	185,568	183,810	172,023	171,365	163,920

Harmonized Data File Examples

- Example 2: More complex harmonized variable.

  - Harmonized Variable CIGTYPE: Type of cigarette

-9: No response	1: Menthol
-3: Refused	2: Non-menthol
-2: Don't know	3: No usual type
-1: Not in universe	

  - Every day and some day smokers, 2003 and later

  - Supplement questions:

    - 2003, 2006-07: Is your usual cigarette brand menthol or non-menthol?

    - 2010-11, 2014-15: Do you usually smoke menthol or non-menthol cigarettes?

  - Two separate questions

    - B2 (Every day smokers), C2 (Some day smokers)

Harmonized Data File Examples

- Example 2:

  - SAS program code to generate unweighted table of CIGTYPE by survey wave

```
Data Harmon;
   Set MyData.Harmon;
   Keep SurWave CigType SRWeight;
Run;

Proc Tabulate Data=Harmon Missing;
   Title1 "Tobacco Use Supplement to The Current Population Survey";
   Title2 "Harmonized Data File";
   Title3 "Variable: CigType";
   Class SurWave CigType;
   Table CigType All="Total",
         SurWave=""*N="Sample"*F=Comma8.;
Run;
```

Harmonized Data File Examples

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: CigType

	1992-1993	1995-1996	1998-1999	2000	2001-2002	2003	2006-2007	2010-2011	2014-2015
	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample
Type of cigarette (menthol/non-menthol, current smokers)									
.	228,552	187,141	176,452	124,582	185,568	.	.	.	.
-9: No response	.	.	.	.	.	110	100	79	59
-3: Refused	.	.	.	.	.	323	190	257	170
-2: Don't know	.	.	.	.	.	107	94	86	64
-1: Not in universe	.	.	.	.	.	149,166	140,522	143,754	140,688
1: Menthol	.	.	.	.	.	8,576	7,718	7,624	6,779
2: Non-menthol	.	.	.	.	.	24,441	22,458	18,758	15,466
3: No usual type	.	.	.	.	.	1,087	941	807	694
Total	228,552	187,141	176,452	124,582	185,568	183,810	172,023	171,365	163,920

Harmonized Data File Examples

- Example 3: Variable constructed from multiple variables.

  - Harmonized Variable WORKBAN: Smoking policy at work

    -9: indeterminate
    -1: not in universe
     1: Not allowed in any indoor public or work areas
     2: Allowed in some/all indoor public or work areas or not applicable
     3: No policy

  - Not available in January and May 2000

  - Constructed from:

    - WKPOLICY: Does your place of work have an official policy that restricts smoking in any way?

    - WKPOLIND: Smoking policy for indoor public or common areas such as lobbies, rest rooms and lunch rooms.

    - WKPOLWRK: Smoking policy for work areas

Harmonized Data File Examples

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: WKPOLICY
Unweighted Counts

| | 1992-1993 | 1995-1996 | 1998-1999 | 2001-2002 | 2003 | 2006-2007 | 2010-2011 | 2014-2015 |
	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample
Does your place of work have an official policy that restricts smoking in any way?								
-9: No response	2,469	548	712	981	20	61	50	56
-3: Refused	.	197	255	338	80	105	106	89
-2: Don't know	1,757	534	594	760	542	662	265	285
-1: Not in universe	124,968	100,825	93,194	98,072	105,190	98,378	96,448	94,071
1: Yes	81,178	73,223	71,378	74,323	70,993	65,279	70,085	65,016
2: No	18,180	11,814	10,319	11,094	6,985	7,538	4,411	4,403
Total	228,552	187,141	176,452	185,568	183,810	172,023	171,365	163,920

Harmonized Data File Examples

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: WKPOLIND
Unweighted Counts

	1992-1993	1995-1996	1998-1999	2001-2002	2003	2006-2007	2010-2011	2014-2015
	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample
Smoking policy for indoor public or common areas such as lobbies, rest rooms and lunch rooms								
-9: No response	405	5	9	8	2	9	17	13
-3: Refused	.	17	10	14	19	21	46	42
-2: Don't know (not read)	412	139	109	137	105	113	138	152
-1: Not in universe	147,374	113,918	105,074	111,245	112,817	106,744	101,280	98,904
1: Not allowed in any public areas	49,661	56,847	58,856	62,679	62,020	56,828	62,468	56,579
2: Allowed in some public areas	28,035	14,288	10,864	10,091	7,916	7,147	6,336	6,540
3: Allowed in all public areas	2,064	1,668	1,378	1,203	765	1,004	687	1,214
4: Not applicable (not read)	601	259	152	191	166	157	393	476
Total	228,552	187,141	176,452	185,568	183,810	172,023	171,365	163,920

Harmonized Data File Examples

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: WKPOLWRK
Unweighted Counts

	1992-1993	1995-1996	1998-1999	2001-2002	2003	2006-2007	2010-2011	2014-2015
	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample
Smoking policy for work areas								
-9: No response	479	9	13	11	6	14	20	18
-3: Refused	..	21	14	18	19	26	50	44
-2: Don't know (not read)	232	76	83	102	90	75	81	79
-1: Not in universe	147,374	113,918	105,074	111,245	112,817	106,744	101,280	98,904
1: Not allowed in any work areas	66,034	64,351	64,525	67,910	66,870	61,192	67,477	62,045
2: Allowed in some work areas	13,193	7,314	5,469	5,128	3,436	3,061	1,841	1,815
3: Allowed in all work areas	1,091	1,336	1,172	1,021	482	815	349	723
4: Not applicable (not read)	149	116	102	133	90	96	267	292
Total	228,552	187,141	176,452	185,568	183,810	172,023	171,365	163,920

Harmonized Data File Examples

WORKBAN	WKPOLICY	WKPOLIND	WKPOLWRK
-9: Indeterminate	-9: No response -3: Refused -2: Don't know	-9: No response -3: Refused -2: Don't know	-9: No response -3: Refused -2: Don't know
1: Not allowed in any public or work areas	1: Yes	1: Not allowed in any public areas	1: Not allowed in any work areas
2: Allowed in some/all indoor public or work areas or not applicable	1: Yes	2: Allowed in some public areas 3: Allowed in all public areas 4: Not applicable	2: Allowed in some work areas 3: Allowed in all work areas 4: Not applicable
3: No policy	2: No	-1: Not in universe	-1: Not in universe

WKPOLICY: Does your place of work have an official policy that restricts smoking in any way
WKPOLIND: Smoking policy for indoor public or common areas such as lobbies, rest rooms and lunch rooms.
WKPOLWRK: Smoking policy for work areas

Harmonized Data File Examples

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: WORKBAN
Unweighted Counts

	1992-1993	1995-1996	1998-1999	2001-2002	2003	2006-2007	2010-2011	2014-2015
	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample
Smoking policy at work (constructed)								
-9: Indeterminate	5,279	1,480	1,737	2,286	816	1,005	665	681
-1: Not in universe	124,968	100,825	93,194	98,072	105,190	98,378	96,448	94,071
1: Not allowed in any indoor public or work areas	45,898	54,560	57,006	61,045	60,799	55,720	61,724	55,884
2: Allowed in some/all indoor public or work areas or not applicable	34,227	18,462	14,196	13,071	10,020	9,382	8,117	8,881
3: No policy	18,180	11,814	10,319	11,094	6,985	7,538	4,411	4,403
Total	228,552	187,141	176,452	185,568	183,810	172,023	171,365	163,920

Hands-on Exercises

NIH NATIONAL CANCER INSTITUTE

Hands-on Exercises

- Exercise 1: Creating a SAS dataset from the ASCII text file

- Exercise 2: Weights

  - Current cigarette smoking status by survey: CIGSTAT * SurWave

- Exercise 3:

  - Home smoking rules by survey and geographic region: Region * SurWave * HMSMKPOL

- Exercise 4:

  - Ever cigar use by survey and sex: SurWave * Sex * EVERCIGR

Hands-on Exercises

- Exercise 1: Creating a SAS dataset from the ASCII text file

  - Download the data and SAS code from TUS-CPS website.

    - Data: harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.zip

    - Main SAS program: harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.sas

    - Formats: harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.formats.sas

  - Unzip the data file

  - Modify <directory> in the main SAS program to reflect where the format program and data are.

    - Filename Harmon "<directory>harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.dat";

    - Libname MyLib "<directory>";

    - %Include "<directory>harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.formats.sas“;

  - Run SAS program to permanent SAS dataset.

Hands-on Exercises

- Exercise 2: Working with Weights

  - Self-response weight needs divided by the number of surveys in the survey wave.

    - There were 2 surveys in 2000.

    - All other survey waves had 3.

  - Generate tables of current cigarette status by survey wave

    - Table of weighted counts

    - Table of weighted percentages

Hands-on Exercises

```sas
LibName  MyData "<directory>";
%Include "<directory>\harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.formats.sas";

Data Harmon;
   Set MyData.Harmon;
   If SurWave=4 Then SRWeight=SRWeight/2;
   Else SRWeight=SRWeight/3;
   Keep SurWave CIGSTAT SRWeight;
Run;

Proc Tabulate Data=Harmon Missing;
   Title1 "Tobacco Use Supplement to The Current Population Survey";
   Title2 "Harmonized Data File";
   Title3 "Variable: CIGSTAT";
   Title4 "Weighted Counts";
   Class SurWave CIGSTAT;
   Var SRWeight;
   Table CIGSTAT All="Total",
         SurWave=""*SRWeight=""*Sum="Population"*F=Comma12.;
Run;

Proc Tabulate Data=Harmon Missing;
   Title4 "Weighted Percentages";
   Class SurWave CIGSTAT;
   Var SRWeight;
   Table CIGSTAT All="Total",
         SurWave=""*SRWeight=""*PctSum<CIGSTAT All>="Percent"*F=9.2;
Run;
```

Hands-on Exercises

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: CIGSTAT
Weighted Counts

	1992-1993	1995-1996	1998-1999	2000	2001-2002	2003	2006-2007	2010-2011	2014-2015
	Population	Population	Population	Population	Population	Population	Population	Population	Population
Current cigarette smoking status									
-9: Indeterminate	691,028	468,256	662,677	615,466	732,705	912,321	977,951	1,185,689	1,255,569
1: Never	96,916,000	102,859,579	109,515,348	111,861,887	117,157,408	129,048,509	135,038,811	149,368,484	162,113,749
2: Every day	37,301,984	37,046,020	35,359,907	34,713,580	34,142,310	32,227,027	32,494,777	28,933,246	25,474,113
3: Some days	8,241,594	8,065,711	8,280,345	9,116,017	8,568,705	7,740,653	8,111,173	7,817,083	7,453,035
4: Former	43,699,325	43,618,104	44,670,838	44,798,956	43,948,725	42,486,585	44,008,921	42,487,311	44,824,090
Total	186,849,931	192,057,670	198,489,114	201,105,905	204,549,853	212,415,094	220,631,634	229,791,812	241,120,556

Hands-on Exercises

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: CIGSTAT
Weighted Percentages

	1992-1993	1995-1996	1998-1999	2000	2001-2002	2003	2006-2007	2010-2011	2014-2015
	Percent	Percent	Percent	Percent	Percent	Percent	Percent	Percent	Percent
Current cigarette smoking status									
-9: Indeterminate	0.37	0.24	0.33	0.31	0.36	0.43	0.44	0.52	0.52
1: Never	51.87	53.56	55.17	55.62	57.28	60.75	61.21	65.00	67.23
2: Every day	19.96	19.29	17.81	17.26	16.69	15.17	14.73	12.59	10.56
3: Some days	4.41	4.20	4.17	4.53	4.19	3.64	3.68	3.40	3.09
4: Former	23.39	22.71	22.51	22.28	21.49	20.00	19.95	18.49	18.59
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Hands-on Exercises

- Exercise 3:

  - Home smoking rules by survey and geographic region: Region * SurWave * HMSMKPOL

  - Not available in January and May 2000

  - Generate tables showing weighed percentages, sample counts and population.

Hands-on Exercises

```
LibName  MyData "<directory>";
%Include "<directory>\harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.formats.sas";

Data Harmon;
  Set MyData.Harmon;
  If SurWave=4 Then SRWeight=SRWeight/2;
  Else SRWeight=SRWeight/3;
  Keep SurWave Region HMSMKPOL SRWeight;
Run;

Proc Tabulate Data=Harmon Missing;
  Title1 "Tobacco Use Supplement to The Current Population Survey";
  Title2 "Harmonized Data File";
  Title3 "Variable: HMSMKPOL";
  Where SurWave In(1,2,3);
  Class SurWave Region HMSMKPOL;
  Var SRWeight;
  Table SurWave=""*(All="Total" Region=" "),
        HMSMKPOL*(SRWeight=""*PctSum<HMSMKPOL>="Percent"*F=7.2)
        All="Total"*(N="Sample"*F=Comma8. SRWeight=""*Sum="Population"*F=Comma12.);
Run:
```

Hands-on Exercises

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: HMSMKPOL

		Home smoking rules						Total	
		-9: No response	-3: Refused	-2: Don't know	1: No one is allowed to smoke anywhere	2: Smoking is allowed in some places or at some times	3: Smoking is allowed anywhere		
		Percent	Percent	Percent	Percent	Percent	Percent	Sample	Population
1992-1993	Total	0.70	.	.	42.79	25.82	30.69	228,552	186,849,931
	1: Northeast	0.77	.	.	40.61	29.31	29.31	53,254	38,272,711
	2: Midwest	0.70	.	.	35.96	28.57	34.77	58,533	43,981,132
	3: South	0.64	.	.	40.69	24.31	34.35	68,273	64,560,523
	4: West	0.72	.	.	55.78	21.89	21.60	48,492	40,035,564
1995-1996	Total	0.31	0.30	0.29	52.82	22.03	24.24	187,141	192,057,670
	1: Northeast	0.44	0.42	0.41	49.91	25.50	23.33	41,039	38,330,257
	2: Midwest	0.24	0.26	0.25	45.75	24.88	28.61	46,527	44,762,497
	3: South	0.25	0.30	0.24	51.48	20.79	26.94	57,107	67,328,833
	4: West	0.36	0.26	0.30	65.27	17.80	16.02	42,468	41,636,084
1998-1999	Total	0.44	0.42	0.41	60.31	19.24	19.19	176,452	198,489,114
	1: Northeast	0.65	0.72	0.60	58.24	21.20	18.59	36,322	38,574,306
	2: Midwest	0.36	0.31	0.25	53.85	21.84	23.39	42,983	45,816,558
	3: South	0.42	0.32	0.38	59.01	18.69	21.18	54,314	70,157,550
	4: West	0.37	0.41	0.43	70.95	15.69	12.16	42,833	43,940,700

Hands-on Exercises

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: HMSMKPOL

| | | Home smoking rules | | | | | | Total | |
| | | -9: No response | -3: Refused | -2: Don't know | 1: No one is allowed to smoke anywhere | 2: Smoking is allowed in some places or at some times | 3: Smoking is allowed anywhere | | |
		Percent	Percent	Percent	Percent	Percent	Percent	Sample	Population
2001-2002	Total	0.38	0.51	0.51	66.28	17.11	15.21	185,568	204,549,853
	1: Northeast	0.47	0.91	0.79	64.52	18.81	14.49	38,906	39,192,663
	2: Midwest	0.32	0.45	0.47	60.30	19.48	18.97	48,140	46,584,508
	3: South	0.36	0.41	0.43	65.07	16.52	17.20	54,026	72,727,315
	4: West	0.37	0.40	0.44	75.72	14.19	8.88	44,496	46,045,367
2003	Total	0.67	0.72	0.42	72.80	12.31	13.08	183,810	212,415,094
	1: Northeast	0.85	0.97	0.54	70.94	14.10	12.59	38,523	41,129,399
	2: Midwest	0.49	0.80	0.50	65.69	15.68	16.84	47,266	48,174,329
	3: South	0.69	0.56	0.36	72.29	11.92	14.17	53,081	75,495,450
	4: West	0.68	0.66	0.34	82.38	7.98	7.95	44,940	47,615,917
2006-2007	Total	0.66	0.66	0.37	77.80	10.35	10.17	172,023	220,631,634
	1: Northeast	0.91	0.74	0.36	76.08	12.47	9.44	33,491	41,259,376
	2: Midwest	0.54	0.69	0.37	72.33	12.80	13.26	42,090	49,150,326
	3: South	0.64	0.61	0.36	76.87	10.18	11.34	54,426	79,823,561
	4: West	0.60	0.64	0.38	86.00	6.49	5.89	42,016	50,398,371

Hands-on Exercises

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: HMSMKPOL

		Home smoking rules						Total	
		-9: No response	-3: Refused	-2: Don't know	1: No one is allowed to smoke anywhere	2: Smoking is allowed in some places or at some times	3: Smoking is allowed anywhere		
		Percent	Percent	Percent	Percent	Percent	Percent	Sample	Population
2010-2011	Total	0.96	0.91	0.28	82.13	8.01	7.71	171,365	229,791,812
	1: Northeast	0.92	0.91	0.24	81.18	9.18	7.58	34,631	42,292,386
	2: Midwest	0.86	1.24	0.36	77.18	10.16	10.20	41,654	50,045,734
	3: South	0.98	0.80	0.24	81.21	8.20	8.57	54,251	84,215,527
	4: West	1.05	0.78	0.29	89.01	4.75	4.12	40,829	53,238,165
2014-2015	Total	0.89	0.97	0.21	84.70	6.18	7.05	163,920	241,120,556
	1: Northeast	1.02	1.10	0.23	83.86	6.95	6.84	28,370	43,495,882
	2: Midwest	0.76	1.20	0.26	81.68	7.61	8.50	35,894	51,111,980
	3: South	0.81	0.81	0.17	83.84	6.53	7.83	58,602	89,905,819
	4: West	1.02	0.89	0.22	89.44	3.75	4.66	41,054	56,606,875

Hands-on Exercises

- Exercise 4:

  - Ever used cigars by sex and survey wave: SurWave * Sex * CIGREVER

  - Create a SurWave recode variable that combines September 1998 data with January and May 2000

  - Generate tables showing weighed percentages, sample counts and population.

Hands-on Exercises

```sas
LibName  MyData "<directory>";
%Include "<directory>\harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.formats.sas";

Proc Format;
  Value SurWavRF
    1 = "1992-1993"
    2 = "1995-1996"
    3 = "1999"
    4 = "1998,2000"
    5 = "2001-2002"
    6 = "2003"
    7 = "2006-2007"
    8 = "2010-2011"
    9 = "2014-2015"
    ;

Data Harmon;
  Set MyData.Harmon;
  If SurYear=1998 Then SurWaveR=4;
  Else SurWaveR=SurWave;
  If SurWave=3 Then SRWeight=SRWeight/2;
  Else SRWeight=SRWeight/3;
  Attrib SurWaveR Label = "Survey wave recode" Format=SurWavRF.;
  Keep SurYear SurMonth SurWave SurWaveR SEX CIGREVER PIPEEVER SRWeight;
Run;

Proc Tabulate Data=Harmon Missing;
  Title1 "Tobacco Use Supplement to The Current Population Survey";
  Title2 "Harmonized Data File";
  Title3 "Variable: CIGREVER (Ever used cigars <regular cigar, cigarillos or little filtered cigars>?";
  TItle4 "No Response, Refused and Don't Know Excluded";
  Where CIGREVER Not in (-9,-3,-2);
  Class SurWaveR Sex CIGREVER;
  Var SRWeight;
  Table SurWaveR="",
        (All="Total" Sex=" ")*(CIGREVER=""*(SRWeight=""*PctSum<CIGREVER>="Percent"*F=7.2)
        All="Total"*(N="Sample"*F=Comma8. SRWeight=""*Sum="Population"*F=Comma12.));
Run;
```

Hands-on Exercises

Tobacco Use Supplement to The Current Population Survey
Harmonized Data File
Variable: CIGREVER (Ever used cigars <regular cigar, cigarillos or little filtered cigars>?
No Response, Refused and Don't Know Excluded

	Total				1: Male				2: Female			
	1: Yes	2: No	Total		1: Yes	2: No	Total		1: Yes	2: No	Total	
	Percent	Percent	Sample	Population	Percent	Percent	Sample	Population	Percent	Percent	Sample	Population
1992-1993	3.96	96.04	227,158	185,726,105	7.97	92.03	97,195	88,643,355	0.30	99.70	129,963	97,082,750
1995-1996	3.64	96.36	186,448	191,310,283	7.31	92.69	78,650	91,284,505	0.28	99.72	107,798	100,025,778
1999	3.75	96.25	113,895	197,958,668	7.37	92.63	49,037	94,550,565	0.45	99.55	64,858	103,408,103
1998,2000	10.62	89.38	185,970	231,956,057	19.78	80.22	80,652	111,006,247	2.20	97.80	105,318	120,949,810
2001-2002	13.40	86.60	184,562	203,403,835	23.50	76.50	80,614	97,345,577	4.14	95.86	103,948	106,058,258
2003	11.72	88.28	181,659	209,848,309	20.46	79.54	78,813	100,429,209	3.71	96.29	102,846	109,419,100
2006-2007	22.64	77.36	170,264	218,264,724	37.20	62.80	73,832	105,032,779	9.13	90.87	96,432	113,231,944
2010-2011	16.48	83.52	168,764	226,527,076	27.03	72.97	74,341	109,021,809	6.68	93.32	94,423	117,505,267
2014-2015	17.27	82.73	161,867	238,103,106	28.21	71.79	72,266	114,248,135	7.17	92.83	89,601	123,854,971

Contacts

- For general TUS questions: ncidccpsbrpadvances@mail.nih.gov

- For TUS harmonized data questions: Email Carolyn Reyes-Guzman (carolyn.reyes-guzman@nih.gov)

- For weighting or TUS variable-specific questions: Email Todd Gibson (gibsont@imsweb.com) and Dr. Reyes-Guzman (carolyn.reyes-guzman@nih.gov)

TUS-CPS website

https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/

www.cancer.gov   **www.cancer.gov/espanol**