HHS Public Access

Author manuscript

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2019 February 11.

Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2016 October ; 25(10): 1402–1410. doi: 10.1158/1055-9965.EPI-16-0244.

Multilevel Small-Area Estimation of Multiple Cigarette Smoking Status Categories Using the 2012 Behavioral Risk Factor Surveillance System

Zahava Berkowitz, **Xingyou Zhang**, **Thomas B. Richards**, **Lucy Peipins**, **S. Jane Henley**, and **James Holt**

National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention (CDC), Chamblee, Georgia.

Abstract

Background: Smoking is the leading preventable cause of death; however, small-area estimates for detailed smoking status are limited. We developed multilevel small-area estimate mixed models to generate county-level estimates for six smoking status categories: current, some days, every day, former, ever, and never.

Method: Using 2012 Behavioral Risk Factor Surveillance System (BRFSS) data (our sample size = 405,233 persons), we constructed and fitted a series of multilevel logistic regression models and applied them to the U.S. Census population to generate county-level prevalence estimates. We mapped the estimates by sex and aggregated them into state and national estimates. We conducted comparisons for internal consistency with BRFSS states' estimates using Pearson correlation coefficients, and external validation with the 2012 National Health Interview Survey current smoking prevalence.

Dr. Zhang was employed by CDC during the time of the project. He is currently employed by the Social, Economic, and Housing Statistics Division, U.S. Census Bureau. **Corresponding Author:** Zahava Berkowitz, Centers for Disease Control and Prevention (CDC), 4770 Buford Highway, Mailstop F-76, Chamblee, GA 30341. Phone: 770-488-4881; Fax: 770-488-3040; zab3@cdc.gov.

Authors' Contributions

Conception and design: Z. Berkowitz, X. Zhang

Development of methodology: Z. Berkowitz, X. Zhang

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Z. Berkowitz, X. Zhang, T.B. Richards, J. Holt

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Z. Berkowitz, X. Zhang, T.B. Richards, L. Peipins, S.J. Henley

Writing, review, and/or revision of the manuscript: Z. Berkowitz, X. Zhang, T.B. Richards, L. Peipins, S.J. Henley, J. Holt

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Z. Berkowitz, X. Zhang

Study supervision: Z. Berkowitz, X. Zhang

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Publisher's Disclaimer: Disclaimer

Publisher's Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the CDC.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Results: Correlation coefficients ranged from 0.908 to 0.982, indicating high internal consistency. External validation indicated complete agreement (prevalence = 18.06%). We found large variations in current and former smoking status between and within states and by sex. County prevalence of former smokers was highest among men in the Northeast, North, and West. Utah consistently had the lowest smoking prevalence.

Conclusions: Our models, which include demographic and geographic characteristics, provide reliable estimates that can be applied to multiple category outcomes and any demographic group. County and state estimates may help understand the variation in smoking prevalence in the United States and provide information for control and prevention.

Impact: Detailed county and state smoking category estimates can help identify areas in need of tobacco control and prevention and potentially allow planning for health care.

Introduction

Cigarette smoking is the leading preventable cause of premature death in the United States. Studies have shown causal association between cigarette smoking and multiple cancers, chronic obstructive pulmonary disease, cardiovascular disease, immune and autoimmune disorders, and nicotine addiction among former and current smokers. Exposure to secondhand tobacco smoke has also been causally linked to cancer, respiratory, and cardiovascular diseases, and to adverse effects on the health of infants and children (1). During 2012 alone, more than 480,000 Americans died of diseases related to cigarette smoking. The associated cost of direct medical care and lost productivity has been estimated to exceed $289 billion (1). The U.S. Preventive Services Task Force has issued recommendations for clinicians to ask all adults about smoking, with the aim of providing smokers with behavioral and pharmacologic treatments for smoking cessation (2). The 2014 Surgeon General's report on the health consequences of smoking described a substantial decline in cigarette smoking among U.S. adults since 1965 (1). Nevertheless, cigarette smoking prevalence in 2014 was estimated at 16.8% (3), translating to 40 million people.

National health surveys have provided reliable estimates of smoking prevalence for large geographic areas such as the entire United States or the states (4–6). Studies using population surveys show considerable variation in health outcomes among states (7). The NCI has estimated county-level prevalence of current and ever smoking for two aggregated periods, 1997–1999 and 2000–2003, by combining data from the Behavioral Risk Factor Surveillance System (BRFSS) and National Health Interview Surveys (NHIS; ref. 8). Similarly, researchers at the University of Washington have used only BRFSS data and estimated county smoking prevalence for men and women for each year from 1996 to 2012 to show that the decline in current cigarette smoking occurred mainly in higher-income counties with larger populations (9).

The prevalence of current cigarette smoking is routinely reported by the U.S. Centers for Disease Control and Prevention (CDC) using BRFSS data (10). However, small-area estimates (SAE) for detailed smoking status, such as an every day smoker, some days smoker, or former smoker, are limited.

The purpose of our study is to extend a SAE method based on a statistical model developed by Zhang and colleagues (11) to generate county estimates for current smokers, former smokers, and never smokers, and to further estimate three additional levels of smoking: ever smoker (former and current), current every day smoker, and current some days smoker. To our knowledge, no other study has estimated cigarette smoking status in such detail.

Materials and Methods

CDC administers BRFSS, a state-based, random-digit-dial survey, annually in collaboration with health departments in the states and the District of Columbia to generate reliable direct estimates. Trained interviewers in each state collect demographic and health-related information on noninstitutionalized adults aged 18 years and older through landline or cell phone interviews. The combined landline and cell phone response rates to the 2012 BRFSS survey ranged from 27.7% in California to 60.4% in South Dakota, with a median rate of 45.2%. Detailed interpretation of the response rate can be found on the BRFSS website (12). To improve the information about the sampled population, we poststratified the BRFSS data with the U.S. Census 2010 population counts (13), which provide the most current detailed information about a community's population, by age, sex, race, and Hispanic or Latino origin groups. County-level poverty rates (≤150% of the federal poverty rate), which are a strong predictor of smoking (14, 15), were extracted from the American Community Survey 5-year county estimates (2007–2011; ref. 16).

We created six smoking status categories from two BRFSS questions as our outcomes: "Have you smoked at least 100 cigarettes in your entire life?" If a responder answered "yes," he or she was further asked, "Do you now smoke cigarettes every day, some days, or not at all?" We categorized those who answered "no" to the first question as "never smoker." Those who answered "every day" or "some days" to the second question were categorized as "every day smoker" or "some days smoker," respectively; these two categories combined are the "current smoker" category. Those who responded "not at all" were categorized as "former smoker." The final category—"ever smoker"—is a combination of "every day," "some days," and "former" smokers.

Additional information for each respondent in the BRFSS survey were age (13 age groups: 18–24, 25–29 … 75–79, or ≥80 years, sex (male or female), and race/ethnicity [eight non-overlapping groups: non-Hispanic (NH) white, NH black, NH American Indian or Alaska Native, NH Asian, NH Native Hawaiian or other Pacific Islander, NH other single race, NH 2 or more races, or Hispanic], as well as the respective county (n = 3,143 counties) and state (n = 51; 50 states and the District of Columbia) identifiers.

Data analysis

To estimate the expected probability of individual smoking status in the United States, in each county nested in a state, we used the 2012 BRFSS data to construct a series of three multilevel logistic mixed models with the whole population, the population of ever smokers, and the population of current smokers. The second and third models progressively included a subset of the previous model's population. Each model included both individual-level fixed

effects (age, sex, race/ethnicity), and county-level poverty, and county- and state-level random effects.

The results from each of these three models included parameters for 13 age categories, 2 sex categories, 8 race/ethnicity categories, county-level poverty, and county- and state-level random effects. We then defined for any county-level i with a missing estimate a county-level random effect μ^{i}_{c} by spatially smoothing its adjacent counties' random effects μ^{j}_{c} ($j \neq i$) and averaging them (17). These newly created county random effects were linked back to the county random effect list.

The estimated parameters from the three models and the updated county random effects were applied in three newly constructed Monte Carlo simulation programs, for which we randomly drew 1,000 samples for each of the parameters and their standard errors, to predict the individual-level expected probability of each smoking category (see model specification below). The county-level random effects in these simulation programs represented the county contextual effects on the outcome rather than one factor for the whole county.

Our multilevel logistic regression models for the prevalence of the six smoking status categories followed the general formula of generalized linear mixed models as follows (11):

Let $P_{ijkcs}(Y_{ijkcs})$ be the probability of an individual having a smoking status assumed to be associated with three level-related factors—individual, county, and state—via the logit link function:

$$P_{ijkcs}\left(Y_{ijkcs}{}^{1}=1\right)=\ \text{logit}^{-1}\left(\alpha_{i}+\beta_{j}+\gamma_{k}+x'_{c}\eta+\mu_{c}+\nu_{s}+e_{ijkcs}\right).$$

$Y_{ijkcs}{}^{1}$ is the self-reported smoking status (1 yes, 0 = no) for an individual in age group i, i = 1 to 13, sex group j, j = 1,2, and race/ethnicity group k, k = 1 to 8 from county c in state s; α_i, β_j, and γ_k are the regression coefficients of age group i, sex group j, and race/ethnicity group k, respectively. x_c is a vector of county-level covariates, and η is a vector of their respective regression coefficients. The prediction model included a product of the county-level poverty status x'_c and the regression coefficient η of the county-level poverty status. μ_c, ν_s, and e_{ijkcs} are the county-level, state-level, and individual residual random effects, which were assumed to be independent and normally distributed.

Model 1 included the entire population and was defined as the general model above with self-reported smoking status $Y_{ijkcs}{}^{1}$ [1 = ever smoker (former, current every day, or current some days), 0 = never smoker]. Model 2 included the population of ever smokers [1 = current (every day and some days), 0 = former smoker], and model 3 included the population of current smokers (1 = every day smoker, 0 = some days smoker). Additional details of the models are presented in Supplementary Methods and Materials.

Aggregating the results for the county-level SAEs into larger units of geography allowed us to estimate prevalence of each smoking status category for each state and the entire United States.

We calculated summary statistics (mean, median, first and third quartiles, interquartile range, minimum, maximum, and range) for the model-based county distributions for each smoking status category, for its total population and by sex, using the univariate procedure. In addition, we mapped the results for four of the model-based county smoking distributions (every day, some days, former, and never) to provide insights into geographic patterns. We conducted external validation of our estimates by comparing our U.S. current smoking estimates with the 2012 NHIS estimates (18). We also evaluated the internal consistency between states' model estimates and the corresponding direct BRFSS estimates with Pearson correlation coefficients. Summary statistics for states' SAE and direct BRFSS estimates were calculated with the MEANS procedure.

The BRFSS multilevel models were fitted with the SAS GLIMMIX procedure (SAS Institute, Inc.). The multilevel simulation models were performed with SAS Ver. 9.3. The calculation of the BRFSS states' summary estimates for internal consistency, as well as of the NHIS estimates for external validation, was performed with SAS-callable SUDAAN (Research Triangle Institute, Research Triangle Park, NC).

Results

The 2012 BRFSS survey included information from a sample of 475,687 adults, with state-level sample sizes ranging from 4,390 in Alaska to 21,895 in Massachusetts. After excluding missing records for age, race, smoking, and county-level poverty rate, the total sample size for our analysis was 405,223. Our post-stratification included U.S. Census 2010 population data from all 3,143 U.S. counties.

Model-based SAEs and external validation

The model-based SAE for overall 2012 current smoking prevalence in the United States was 18.06% (Table 1), the same as the prevalence estimate reported by NHIS for that year. The direct BRFSS estimate of current smoking prevalence in 2012 was18.85%.

The model-based SAEs for current smoking prevalence for men and women were 21.20% and 16.41%, respectively, and within the 95% confidence intervals (CI) of the corresponding NHIS prevalence estimates (20.49%, 95% CI, 19.61%–21.39% and15.81%, 95% CI, 15.10%–16.56%). The direct BRFSS 2012 estimates for men and women were 21.32% and 16.52%, respectively.

Results from our model-based SAEs at the county level revealed large variations in the estimated prevalence of current smoking, ranging from 5.76% to 42.02%, with a median of 19.53%. The county-level mean current smoking prevalence was an estimated 22.46% for males and 17.46% for females, with an overall percentage range of 39.04% and 33.74%, respectively. This percentage difference by gender held for every day, some days, and former smokers both for the U.S. overall and county-level means.

Our analysis also showed that, in 2012, the U.S. overall and the county-level mean estimate of never smokers was higher among women than among men (62.56% vs. 50.83% and 60.86% vs.49.10%, respectively). In addition, county-level never smoking prevalence ranged from a minimum of 33.42% for men to a maximum of 87.24% for women, and the median for women was nearly 12% higher than that of men.

More than half of those classified as ever smokers were former smokers. More men than women were former smokers (27.97% vs. 21.04%, respectively, for the U.S. overall and 28.44% vs. 21.69%, respectively, for the counties).

Model-based SAEs versus direct BRFSS estimates for the assessment of internal consistency

The Pearson correlation coefficients between our state-level model-based SAEs for the entire United States and the direct BRFSS survey estimates ranged from 0.908 for some days smokers to 0.982 for former smokers (Table 2), indicating high internal consistency. The correlation coefficients for men ranged from 0.857 for ever smoker (and never smoker, which is the complement to the whole population studied) to 0.919 for every day smoker. Correlation coefficients for women ranged from 0.816 for some days smokers to 0.957 for every day smokers. Except for the correlation for some days smokers, women's correlation coefficients were higher than those of men. Comparison between summary statistics for the states' model-based estimates with those of BRFSS for each smoking category showed that the estimates were similar, with the BRFSS estimates being slightly higher.

Geographic variation in estimated current and former smoking prevalence among the 50 states, DC, and the 3,143 counties

Utah consistently had the lowest prevalence of current smoking overall (9.64%) and for men (11.24%) and women (8.05%; Table 3). The remaining nine states with the lowest prevalence were grouped in the West and Northwest (Hawaii, California, Washington, Oregon, and Idaho), the Northeast (Vermont, Connecticut, and New Hampshire), and Maryland. The highest prevalence of current smoking was observed in part of the Midwest and in the South and, particularly in West Virginia (24.81% overall, 27.79% for men, and 21.96% for women) and Kentucky (24.56% overall, 27.52% for men, and 21.76% for women). The remaining states with the highest prevalence of current smoking among men and women also included Alabama, Tennessee, Missouri, Arkansas, Oklahoma, Indiana, Louisiana, Ohio, and Mississippi. The prevalence among females was consistently lower than that of males.

Consistent with the estimates of state-specific current smoking prevalence, the counties with the lowest prevalence were in Utah and California (<10%, data not shown). The counties with the highest prevalence of current smoking (>33%) were in Arkansas, South Dakota, North Dakota, and Wisconsin, even though the latter three states were not among the those with the highest current smoking prevalence. Additional findings show variation of smoking levels in Arizona, New Mexico, and states in the Northwest and North, and by sex (Fig. 1). The majority of current smokers were every day smokers. Although the county every day smoking prevalence was higher among men than among women, most of the counties where

the prevalence was highest among men were the same as those with the highest prevalence among women. The majority of some days smokers reside in the South, in some of the counties in Alaska and Nevada, and are more often men in states in the North and Northeast (Fig. 1).

By state, 15.30% to 30.62% of men and women were classified as former smokers (Table 4). The highest and lowest estimated prevalence of former smokers included similar states for the entire population, and for both sexes, with few variations. Utah had the lowest estimates: 15.30% for the entire population, 18.26% for men, and 12.38% for women. Other states with the lowest percentages of former smokers were Texas, Mississippi, Georgia, Louisiana, the District of Columbia, Alabama, and Tennessee. Lower prevalence of former smoking was also observed for men in Kansas, for both sexes in Idaho, for the entire population and for women in California, and for the entire population in Indiana. The Northeast, on the other hand, had the highest prevalence of former smokers; Maine, Vermont, and New Hampshire led the states in this region, followed by Connecticut, Rhode Island, and Massachusetts. Other states with a high prevalence of former smokers were Florida, Oregon, and South Dakota, for the entire population and for men and women; Montana for the entire population and for men; and Delaware for women.

Distributions of former smoking prevalence by county reveal that counties with highest prevalence of former smoking in men were similar to the counties with the highest prevalence of former smoking in women, even though prevalence in men was 8% to 9% higher than prevalence in women (Fig. 1). Counties with the lowest prevalence of former smokers were in Utah, Texas, Mississippi, Idaho, and Kansas. The 10 counties with the lowest prevalence of former smokers had less than 15% overall (data not shown). The 10 counties with the highest prevalence of former smokers had more than 34% and were in Florida, Maine, Vermont, Massachusetts, and New Hampshire (data not shown).

Discussion

We presented an extension of a statistical model to generate county estimates for six categories of smoking status. We validated our model-based accuracy by comparing its estimates to the 2012 NHIS estimates for the entire United States and by sex, and assessed its internal consistency by comparing its estimates with the corresponding 2012 BRFSS estimates. Both our external and internal consistency assessments showed the model-based SAEs to be reliable. Our findings were in complete agreement with the NHIS current smoking estimate for the entire United States and very similar to the NHIS estimates for men and women. Our model-based SAEs for prevalence of current smoking were slightly closer to the corresponding estimates of NHIS than those of BRFSS. Our model-based SAEs for prevalence of former smoking were slightly higher than NHIS estimates.

Our findings revealed large variations in the prevalence of current and former cigarette smoking among states and among counties in the states. Although the prevalence for both smoking categories was higher among men than women, smoking distribution patterns were mostly similar and often concentrated in certain areas or regions of the country. Consistent with a previous study (9), the highest estimates of current smoking were in the South and

part of the Midwest, including Appalachian states (West Virginia, Kentucky, Tennessee), where a large percentage of the population is poor with low educational attainment. These states have the lowest excise tax on cigarettes in the United States (19), and some are among the biggest tobacco growers in the United States (20). Other concentrated areas of high-risk populations were observed in Alaska, several other states (Nevada, Arizona, New Mexico), and certain counties within states (South Dakota and North Dakota), with high poverty rates and large populations of American Indians (21, 22). State or federal policies to curb tobacco smoking, such as excise tax, may not apply to Native Americans in their tribal lands unless authorized by federal law (23, 24), which may partly explain the higher tobacco use rates and lower cessation among these vulnerable groups. This finer detail about high levels of smoking prevalence at the county level highlights areas for investigation or for comprehensive tobacco control efforts. An important contribution of our study to the literature and to county and state planning efforts is data about the distribution of former smokers, for whom data have been sparse. A significant proportion of lung cancers arise among former smokers, even years after smoking cessation (25, 26), and lung cancer screening is recommended for some former smokers on the basis of their age, smoking history, and time since quitting (27). Our model-based SAEs for former smokers highlight geographic areas where reductions in cigarette smoking have occurred and areas where additional information (e.g., age, pack-years, and years since cessation) would be helpful to determine eligibility for lung cancer screening. Although our model-based SAEs showed that the largest percentage of former smokers resided in more affluent states in the Northeast, where excise tax is among the highest in the country and cessation or other programs might exist (28), some of these states also had the lowest prevalence of current smoking. However, some counties in this region had very high prevalence of former smoking. Additional examples show that although few counties in South Dakota were among the counties with the highest estimated prevalence of current cigarette smoking, the state as a whole was among the states with the largest percentage of former smokers.

Our study has at least two limitations. First, questions are not currently available on the BRFSS questionnaire to estimate eligibility for lung cancer screening with low-dose computed tomography, including smoking pack-years or whether a person has stopped smoking for at least 15 years. More detailed quantification of smoking status would potentially provide a more useful categorization of current and former smoking. Second, BRFSS data relied on self-reported information, which might have introduced some bias, most likely underestimates.

Using the largest health survey in the United States, BRFSS, together with county-level poverty data, is a study strength. In addition, the models combine prediction from a unit-level (individual-level data as outcomes) with multilevel regression and post stratification that included both geographic and demographic characteristics (29, 30). Further, our small-area estimation method allows better integration with other data available at the county level, such as data from the American Community Survey, for more fine-tuned analysis of factors associated with smoking status. Our modeling design can provide reliable estimates, is flexible, and can be applied to more than two categorical outcomes and to any demographic group. Moreover, our models were based on a method that was internally and externally

validated (31). The county estimates can be aggregated into larger geographic areas, such as states and the entire United States.

Providing county and state estimates of all categories of smoking status across the United States will serve to identify areas where the impact of tobacco control and prevention efforts may be the greatest and allow planning for health care. Understanding the variation in smoking prevalence across and within different geographic units, some of which may have weak tobacco control policies, can provide information for tailoring proven prevention efforts (28, 32) to the appropriate population level and for developing strategies to address interventions at multiple levels. The results of these model-based estimates can also be used for models of various health outcomes and with data sources other than BRFSS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. US Department of Health and Human Services. The health consequences of smoking—50 years of progress: A report of the Surgeon General. Atlanta, GA: US Department of Health and Human Services, CDC; 2014 Available from: http://www.surgeongeneral.gov/library/reports/50-years-of-progress/sgr50-chap-12.pdf.
2. U.S. Department of Health and Human Services. Agency for Healthcare Research and Quality. The Guide to Clinical Preventive Services 2014 Recommendation of the U.S. Preventive Services Task Force 2014. Tobacco Use in Adults, Counseling and Interventions. [Page 61]. AHRQ Pub. No. 14–05158. ISBN 978–1–58763–439–0. 5 2014 Available from: https://www.uspreventiveservicestaskforce.org/Home/GetFileByID/989.
3. Jamal A, Homa DM, O'Connor E, Babb SD, Caraballo RS, Singh T. Current cigarette smoking among adults - United States, 2005–2014. MMWR Morb Mortal Wkly Rep 2015;64:1233–40. [PubMed: 26562061]
4. Henley SJ, Eheman CR, Richardson LC, Plescia M, Asman KJ, Dube SR, et al. State-specific trends in lung cancer incidence and smoking—United States, 1999–2008. MMWR Morb Mortal Wkly Rep 2011; 60:1243–7. [PubMed: 21918494]
5. Jamal A, Agaku IT, O'Connor E, King BA, Kenemer JB, Neff L. Current cigarette smoking among adults—United States, 2005–2013. MMWR Morb Mortal Wkly Rep 2014;63;1108–12. [PubMed: 25426653]
6. Office on Smoking and Health, National Center for Chronic Disease Prevention and Health Promotion. Smoking and tobacco use: data and statistics; state data. [Page last updated: March 8, 2014, last reviewed: February 8, 2016]. Available from: http://www.cdc.gov/tobacco/data_statistics/state_data/.
7. Bryan L, Westmaas L, Alcaraz K, Jemal A. Cigarette smoking and cancer screening underutilization by state: BRFSS 2010. Nicotine Tob Res 2014;16:1183–9. [PubMed: 24696511]
8. Raghunathan ET, Xie D, Schenker N, Parsons VL, Davis WW, Dodd KW, et al. Combining information from two surveys to estimate county-level prevalence rates of cancer risk factors and screening. J Am Stat Assoc 2007;102:474–86.
9. Dwyer-Lindgren L, Mokdad AH, Srebotnjak T, Flaxman AD, Hansen GM, Murray CJL. Cigarette smoking prevalence in US counties: 1996–2012. Popul Health Metr 2014;12:5. [PubMed: 24661401]
10. Office of Surveillance, Epidemiology and Laboratory Services. Centers for Disease Control and Prevention Behavioral Risk Factor Surveillance System: 2012 Summary Data Quality Report. Atlanta, GA: Centers for Disease Control and Prevention; 2014 Available from: http://www.cdc.gov/brfss/annual_data/2012/pdf/summarydataqualityreport2012_20130712.pdf.

11. Zhang X, Holt JB, Lu H, Wheaton AG, Ford ES, Greenlund KJ, et al. Multilevel regression and post stratification for small-area estimation population health outcomes: A case study of chronic obstructive pulmonary disease prevalence using the Behavioral Risk Factor Surveillance System. Am J Epidemiol 2014;179:1025–33. [PubMed: 24598867]
12. Office of Surveillance, Epidemiology and Laboratory Services. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System: BRFSS combined Landline and cell phone weighted response rates by state; 2012 Available from: http://www.cdc.gov/brfss/annual_data/2012/pdf/summarydataqualityreport2012_20130712.pdf.
13. United States' Census 2010 2010 Census Data. 2010 Census Data Product Descriptions. [Last updated 2013]. Available from: http://www.census.gov/2010census/data/.
14. Ciapponi A, Bardach A, Glujovsky D, Aruj P, Mazzini A, Linetzky B, et al. Systematic review of the link between tobacco and poverty World Health Organization 2011 Project conducted for WHO by the Institute of Clinical Effectiveness and Health Policy. Argentine Cochrane Centre IECS. IberoAmerican Cochrane Network Available from: http://apps.who.int/iris/bitstream/10665/44453/1/9789241500548_eng.pdf.
15. Scarinci IC, Robinson LA, Alfano CM, Zbikowski SM, Klesges RC. The relationship between socioeconomic status, ethnicity, and cigarette smoking in urban adolescents. Prev Med 2002;34:171–8. [PubMed: 11817912]
16. United States' Census Bureau. American Community Survey (ACS). [2012 data release. Updated October 23, 2013]. Available from: http://www.census.gov/program-survey/acs/.
17. US Census Bureau, US Department of Commerce. Standard hierarchy of census geographic entities. Washington, DC: US Census Bureau; 2010 Available from: http://www2.census.gov/geo/pdfs/reference/geodiagram.pdf.
18. Agaku IT, King BA, Dube SR. Current cigarette smoking among adults -United States, 2005–2012. MMWR Morb Mortal Wkly Rep 2014;63:29–34. [PubMed: 24430098]
19. Lisa Mahapatra. The price of cigarettes: How much does a pack cost in each US state? [MAP]. 2014 Available from: http://www.ibtimes.com/price-cigarettes-how-much-does-pack-cost-each-us-state-map-1553445.
20. U.S. Department of Agriculture. 2012 Census of Agriculture: United States Summary and State Data, Volume 1, Part 51[PDF–34 MB]. Washington: U.S. Department of Agriculture, National Agricultural Statistics; 2012. Available from: https://www.agcensus.usda.gov/Publications/2012/.
21. United States' Census Bureau. Census quick facts. Available from: http://www.census.gov/quickfacts/.
22. Plescia M, Henley SJ, Pate A, Underwood JM, Rhodes K. Lung cancer deaths among American Indians and Alaska Natives, 1990–2009. Am J Public Health 2014;104 Suppl 3:S388–95. [PubMed: 24754613]
23. LaFaive MD, Fleenor P, Nesbit T. Tax-exempt cigarette sales on Indian reservations. Mackinac Center for Public Policy 2008 Available from: https://www.mackinac.org/10038.
24. Chriqui J, DeLong H, Gourdet C, Chaloupka F, Edwards SM, Xu X, et al. Centers for Disease Control and Prevention (CDC). Use of tobacco tax stamps to prevent and reduce illicit tobacco trade–United States, 2014. MMWR Morb Mortal Wkly Rep 2015;64:541–6. [PubMed: 26020136]
25. Mong C, Garon EB, Fuller C, Mahtabifard A, Mirocha J, Mosenifar Z, et al. High prevalence of lung cancer in a surgical cohort of lung cancer patients a decade after smoking cessation. J Cardiothorac Surg 2011;6:19. [PubMed: 21352550]
26. Rostron B Smoking-attributable mortality by cause in the United States: Revising the CDC's data and estimates. Nicotine Tob Res 2013; 15:238–46. [PubMed: 22592446]
27. Moyer VA; US Preventive Services Task Force. Screening for lung cancer: US Preventive Services Task Force recommendation statement. Ann Intern Med 2014;160:330–8. [PubMed: 24378917]
28. Centers for Disease Control and Prevention. Best practices for comprehensive tobacco control programs—2014. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014 Available from: http://www.cdc.gov/tobacco/stateandcommunity/best_practices/.

29. Gelman A, Little TC. Post stratification into many categories using hierarchical logistic regression. Surv Methodol 1997;23:127–35.

30. Park DK, Gelman A, Bafumi J. Bayesian multilevel estimation with post-stratification: State-level estimates from national polls. Polit Anal 2004: 12:375–85.

31. Zhang X, Holt JB, Yun S, Lu H, Greenlund KJ, Croft JB. Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the behavioral risk factor surveillance system. Am J Epidemiol 2015;182:127–37. [PubMed: 25957312]

32. Siu A. Behavioral and pharmacotherapy interventions for Tobacco smoking cessation in adults, including pregnant women: U.S. Preventive Services task Force recommendation statement. Ann Intern Med. 2015 DOI:HYPERLINK “http://dx.doi.org/10.7326/M15-2023 ” 10.7326/M15-2023 . Available from: 10.7326/M15-202310.7326/M15-2023http://www.ncbi.nlm.nih.gov/pubmed/26389730” 10.7326/M15-2023 . Available from: http://www.ncbi.nlm.nih.gov/pubmed/26389730 .

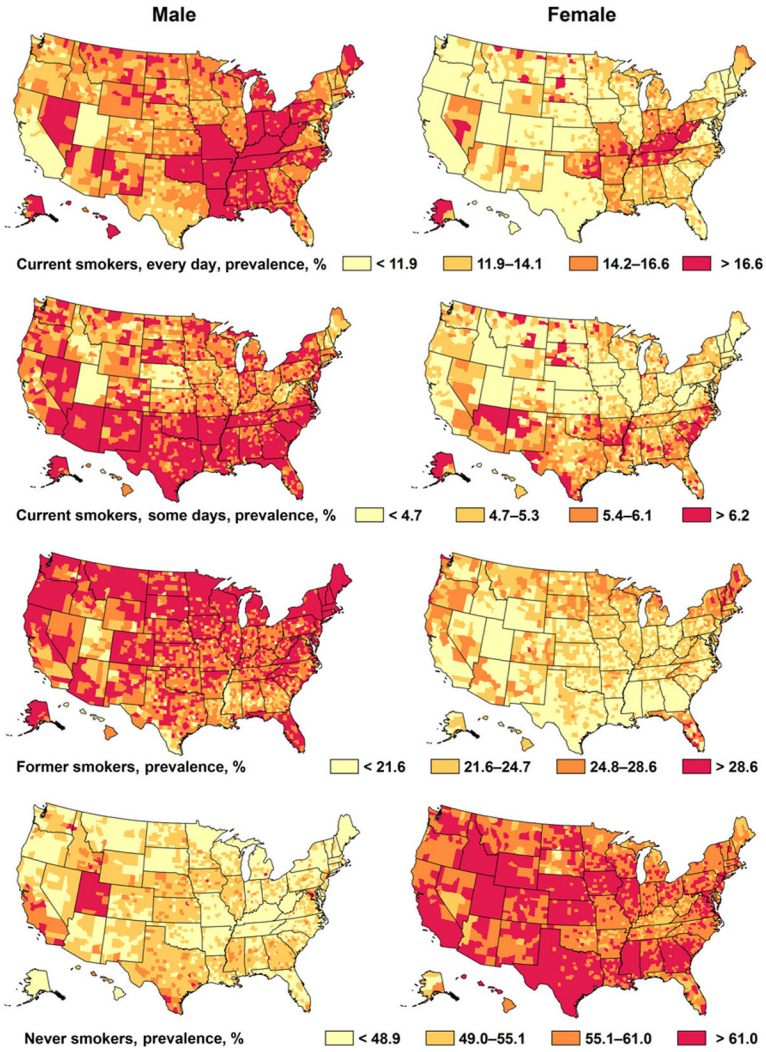

Figure 1.

Model-based estimated prevalence (%) of four smoking levels, by county and gender, United States, 2012. Note: Cut points are quartiles for the combined male and female totals, for each of the four smoking categories.

Table 1.

Model-based smoking levels summary estimates (%) for the total population and the 3,143 counties and by gender, and current smoking total population estimates for NHIS 2012 and BRFSS 2012 to assess external validity and by gender

	Prevalence estimate, %									
Smoking level	**Mean**	**Minimum**	**Lower quartile**	**Median**	**Upper quartile**	**Maximum**	**Inter quartile range**	**Overall range**	**NHIS 2012**	**BRFSS 2012**
Current smoker										
Overall										
U.S.	18.06								18.06	18.85
Counties	19.94	5.76	17.48	19.53	22.18	42.02	4.71	36.26		
Males										
U.S.	21.20								20.49	21.32
Counties	22.46	6.86	19.74	22.06	24.96	45.90	5.22	39.04		
Females										
U.S.	16.41								15.81	16.52
Counties	17.46	4.69	15.20	17.07	19.42	38.44	4.22	33.74		
Every day smoker										
Overall										
U.S.	12.54									
Counties	14.48	4.04	12.39	13.95	16.20	27.05	3.82	23.01		
Males										
U.S.	15.13									
Counties	16.39	4.84	14.10	15.83	18.38	29.63	4.28	24.79		
Females										
U.S.	11.54									
Counties	12.61	3.25	10.71	12.14	14.15	24.66	3.44	21.40		
Some days smoker										
Overall										
U.S.	5.52									
Counties	5.46	1.72	4.75	5.31	5.99	16.31	1.24	14.58		
Males										
U.S.	6.07									

	Prevalence estimate, %									
Smoking level	**Mean**	**Minimum**	**Lower quartile**	**Median**	**Upper quartile**	**Maximum**	**Inter quartile range**	**Overall range**	**NHIS 2012**	**BRFSS 2012**
Counties	6.07	2.02	5.26	5.89	6.67	17.50	1.40	15.49		
Females										
U.S.	4.86									
Counties	4.85	1.44	4.24	4.73	5.32	15.20	1.07	13.76		
Former smoker										
Overall										
U.S.	23.93									
Counties	25.02	9.15	22.93	25.10	27.21	38.39	4.28	29.24		
Males										
U.S.	27.97									
Counties	28.44	11.54	26.09	28.53	30.93	41.68	4.83	30.14		
Females										
U.S.	21.04									
Counties	21.69	7.01	19.82	21.78	23.61	34.80	3.78	27.79		
Ever smoker										
Overall										
U.S.	41.99									
Counties	44.96	16.09	42.30	45.00	48.03	60.71	5.74	44.62		
Males										
U.S.	49.16									
Counties	50.90	19.50	48.15	51.03	54.00	66.58	5.86	47.08		
Females										
U.S.	37.44									
Counties	39.14	12.76	36.55	39.13	42.07	55.27	5.51	42.51		
Never smoker										
Overall										
U.S.	58.01									
Counties	55.04	39.29	51.97	55.00	57.70	83.91	5.74	44.62		
Males										
U.S.	50.83									

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2019 February 11.

	Prevalence estimate, %									
Smoking level	**Mean**	**Minimum**	**Lower quartile**	**Median**	**Upper quartile**	**Maximum**	**Inter quartile range**	**Overall range**	**NHIS 2012**	**BRFSS 2012**
Counties	49.10	33.42	46.00	49.00	51.85	80.50	5.86	47.08		
Females										
U.S.	62.56									
Counties	60.86	44.73	57.93	60.87	63.45	87.24	5.51	42.51		

Abbreviation: U.S., United States.

Table 2.

Pearson correlation coefficient and summary statistics for the states' model-based SAE and direct BRFSS 2012 smoking levels estimates for the total population and by gender

		Prevalence estimate, %						
Smoking level	**ρ^a**	**Minimum**	**Lower quartile**	**Median**	**Upper quartile**	**Maximum**	**Inter quartile range**	**Overall range**
Current smoker								
All	0.947							
Model		9.64	17.08	18.71	20.28	24.81	3.19	15.17
BRFSS12		10.56	17.33	19.62	22.53	28.29	5.20	17.73
Males	0.908							
Model		11.24	19.40	21.25	22.90	27.79	3.49	16.56
BRFSS12		11.92	19.48	21.62	24.57	30.60	5.09	18.67
Females	0.930							
Model		8.05	14.86	16.27	17.86	21.96	3.00	13.91
BRFSS12		9.21	15.66	17.44	19.68	27.61	4.02	18.40
Every day smoker								
All	0.961							
Model		6.75	11.72	12.94	14.36	20.38	2.64	13.63
BRFSS12		7.20	11.92	13.54	15.92	23.98	4.01	16.79
Males	0.919							
Model		7.92	13.31	14.72	16.32	22.91	3.01	14.99
BRFSS12		7.88	13.17	15.04	17.48	25.19	4.31	17.31
Females	0.957							
Model		5.59	10.15	11.27	12.55	17.96	2.40	12.37
BRFSS12		6.23	10.77	12.85	14.32	23.45	3.54	17.22
Some days smoker								
All	0.908							
Model		2.89	5.01	5.48	5.92	7.39	0.90	4.50
BRFSS12		3.36	5.10	5.67	6.25	8.21	1.15	4.85
Males	0.871							
Model		3.32	5.61	6.09	6.62	8.28	1.01	4.96

		Prevalence estimate, %						
Smoking level	**ρ[a]**	**Minimum**	**Lower quartile**	**Median**	**Upper quartile**	**Maximum**	**Inter quartile range**	**Overall range**
BRFSS12		4.05	6.43	6.30	7.11	9.28	1.69	5.23
Females	0.816							
Model		2.46	4.50	4.90	5.31	6.61	0.81	4.15
BRFSS12		2.68	4.60	4.93	5.37	7.27	0.77	4.59
Former smoker								
All	0.982							
Model		15.30	23.22	24.34	25.67	30.62	2.44	15.31
BRFSS12		16.52	23.93	25.24	26.89	31.57	2.95	15.05
Males	0.874							
Model		18.26	26.42	27.83	29.46	34.86	3.03	16.60
BRFSS12		20.85	27.35	28.84	29.85	33.24	2.49	12.39
Females	0.946							
Model		12.38	19.98	20.84	22.17	26.74	2.19	14.36
BRFSS12		12.27	20.40	21.57	23.52	30.26	3.12	17.99
Ever smoker								
All	0.958							
Model		24.94	41.52	43.51	45.58	49.72	4.06	24.78
BRFSS12		27.08	43.01	45.62	47.64	53.75	4.63	26.67
Males	0.857							
Model		29.50	47.60	49.49	51.69	55.88	4.10	26.39
BRFSS12		32.77	48.55	51.01	52.93	60.64	4.39	27.86
Females	0.947							
Model		20.44	35.90	37.81	39.79	43.85	3.89	23.41
BRFSS12		21.48	37.31	40.46	42.58	48.22	5.28	26.74
Never smoker								
All	0.958							
Model		50.28	54.42	56.49	58.48	75.06	4.06	24.78
BRFSS12		46.25	52.36	54.38	56.99	72.92	4.63	26.67
Males	0.857							
Model		44.12	48.31	50.51	52.40	70.50	4.10	26.39

Smoking level	ρ[a]	Prevalence estimate, %						
		Minimum	Lower quartile	Median	Upper quartile	Maximum	Inter quartile range	Overall range
BRFSS12		39.36	47.07	48.99	51.45	67.23	4.39	27.86
Females	0.947							
Model		56.15	60.21	62.19	64.10	79.56	3.89	23.41
BRFSS12		51.78	57.42	59.34	62.69	78.52	5.28	26.74

[a]Pearson correlation coefficient for the correlation between model-based SAEs and BRFSS 2012 by smoking status level for the respective total population and by gender.

Table 3.

The 10 states with the lowest and highest ranking of current cigarette smoking prevalence estimates (%) respectively

	Male and female		Male		Female	
Rank	**State**	**Prevalence (SE)**	**State**	**Prevalence (SE)**	**State**	**Prevalence (SE)**
Lowest ranking						
1	Utah	9.64 (0.11)	Utah	11.24 (0.18)	Utah	8.05 (0.13)
2	California	13.01 (0.06)	California	14.93 (0.10)	California	11.15 (0.07)
3	Hawaii	14.96 (0.17)	Hawaii	17.18 (0.26)	Hawaii	12.77 (0.23)
4	Vermont	15.72 (0.15)	Vermont	17.63 (0.23)	Vermont	13.9 (0.20)
5	Washington	16.06 (0.10)	Washington	18.16 (0.15)	Washington	14.01 (0.12)
6	Maryland	16.11 (0.09)	Connecticut	18.32 (0.23)	Maryland	14.08 (0.12)
7	Connecticut	16.18 (0.15)	Maryland	18.34 (0.14)	Connecticut	14.21 (0.19)
8	Oregon	16.41 (0.12)	New Hampshire	18.50 (0.28)	Oregon	14.39 (0.16)
9	New Hampshire	16.53 (0.18)	Oregon	18.52 (0.19)	Idaho	14.49 (0.17)
10	Idaho	16.68 (0.13)	Idaho	18.90 (0.21)	New Hampshire	14.66 (0.24)
Highest ranking						
42	Mississippi	20.92 (0.07)	Mississippi	23.90 (0.12)	Alabama	18.25 (0.10)
43	Ohio	21.12 (0.08)	Ohio	23.96 (0.13)	Ohio	18.50 (0.10)
44	Louisiana	21.35 (0.08)	Louisiana	24.27 (0.13)	Louisiana	18.62 (0.11)
45	Indiana	22.35 (0.09)	Indiana	25.25 (0.14)	Indiana	19.61 (0.12)
46	Oklahoma	22.62 (0.09)	Oklahoma	25.44 (0.14)	Oklahoma	19.91 (0.12)
47	Arkansas	22.66 (0.10)	Arkansas	25.52 (0.15)	Arkansas	19.96 (0.13)
48	Missouri	22.78 (0.10)	Missouri	25.75 (0.16)	Missouri	20.01 (0.13)
49	Tennessee	23.51 (0.09)	Tennessee	26.56 (0.14)	Tennessee	20.69 (0.12)
50	Kentucky	24.56 (0.09)	Kentucky	27.52 (0.13)	Kentucky	21.76 (0.12)
51	West Virginia	24.81 (0.11)	West Virginia	27.79 (0.18)	West Virginia	21.96 (0.15)

NOTE: Results from model-based SAE, total, and by gender, 2012.

Table 4.

The 10 states with the lowest and highest ranking of former cigarette smoking prevalence estimates (%) respectively

	Male and female		Male		Female	
Rank	**State**	**Prevalence (SE)**	**State**	**Prevalence (SE)**	**State**	**Prevalence (SE)**
Lowest ranking						
1	Utah	15.30 (0.03)	Utah	18.26 (0.04)	Utah	12.38 (0.04)
2	Texas	20.14 (0.01)	Texas	23.35 (0.02)	Texas	17.05 (0.02)
3	Mississippi	20.33 (0.01)	Mississippi	23.49 (0.02)	Mississippi	17.45 (0.02)
4	Georgia	20.97 (0.01)	Louisiana	24.16 (0.02)	Georgia	17.88 (0.02)
5	Louisiana	20.99 (0.02)	Georgia	24.31 (0.02)	DC	17.99 (0.15)
6	DC	21.40 (0.09)	DC	25.32 (0.11)	Louisiana	18.03 (0.02)
7	Alabama	22.08 (0.02)	Alabama	25.41 (0.02)	Idaho	18.97 (0.04)
8	Tennessee	22.32 (0.02)	Tennessee	25.50 (0.02)	Alabama	19.04 (0.02)
9	Indiana	22.49 (0.03)	Idaho	26.05 (0.03)	California	19.13 (0.03)
10	California	22.66 (0.02)	Kansas	26.13 (0.03)	Tennessee	19.39 (0.02)
Highest ranking						
42	Montana	26.46 (0.03)	Montana	30.22 (0.04)	Delaware	22.80 (0.11)
43	South Dakota	26.74 (0.03)	South Dakota	30.30 (0.04)	Oregon	23.04 (0.05)
44	Oregon	26.78 (0.03)	Oregon	30.69 (0.04)	South Dakota	23.24 (0.04)
45	Massachusetts	27.05 (0.03)	Massachusetts	30.89 (0.04)	Massachusetts	23.57 (0.05)
46	Rhode Island	27.45 (0.07)	Rhode Island	31.23 (0.08)	Florida	24 01 (0.04)
47	Florida	27.60 (0.03)	Florida	31.46 (0.03)	Rhode Island	24.03 (0.12)
48	Connecticut	28.28 (0.05)	Connecticut	32.28 (0.06)	Connecticut	24.60 (0.08)
49	New Hampshire	30.50 (0.05)	New Hampshire	34.77 (0.05)	New Hampshire	26.42 (0.07)
50	Vermont	30.58 (0.04)	Maine	34.77 (0.04)	Vermont	26.50 (0.07)
51	Maine	30.62 (0.04)	Vermont	34.86 (0.05)	Maine	26.74 (0.06)

NOTE: Results from model-based SAE, total, and by gender, 2012.