Within-State Geographic Patterns of Health Insurance Coverage and Health Risk Factors in the United States

Linda Williams Pickle, PhD, Yuchen Su, MS

Background: A number of health risk factors have been associated with the incidence and mortality of common diseases. Although knowing risk factor patterns at a small-area level would be useful for ecologic analyses and prevention program planning, risk factor data are generally published only at the state or regional level in the United States. This study presents maps of within-state patterns of several such factors.

Methods: Responses to Behavioral Risk Factor Surveillance System (BRFSS) questions about smoking, obesity, health insurance, and mammography use were aggregated for 1992–1998 by county. These data were then geographically smoothed by adjusting each county's proportional response based on the responses of its neighboring counties.

Results: The maps show risk factor patterns consistent with published state-level maps, but also identify within-state variations masked by aggregation to the larger geographic units.

Conclusions: The risk factor maps presented should permit a better understanding of localized patterns of health risk behaviors and access to health care as well as help to target intervention activities in the U.S. areas that most need them.

Medical Subject Headings (MeSH): epidemiology; health surveys; maps; small-area analysis; statistics, nonparametric (Am J Prev Med 2002;22(2):75–83)

Introduction

A number of personal lifestyle habits have been identified as risk factors for the incidence or mortality of major diseases. For example, cigarette smoking has long been associated with lung cancer,[1] and obesity has been implicated in the initiation of diabetes, increased mortality among men,[2] and is possibly related to the occurrence of coronary heart disease[3] and cancer.[4] Likewise, lack of health insurance can lead to delays in seeking health care and lower utilization of screening or preventive services in some areas.[5–7] This in turn increases the risk that medical conditions that occur will not be diagnosed and treated at an early stage, thus potentially leading to premature death.

Geographic differences in health insurance coverage, physician practices, and health behaviors are known to occur.[5,8–11] Information about these geographic patterns is useful for ecologic studies (e.g., for hypothesis generation relative to disease patterns), and for planning and evaluating risk-reduction activities in high-rate communities. However, health risk data are generally not available for geographic units smaller than states and so are not very useful for the many diseases that vary across communities.[12] The Behavioral Risk Factor Surveillance System (BRFSS) collects and reports health risk data at the state level, but the county of each respondent's residence is also identified. Because sample sizes in any single year are too small to provide reliable county estimates, we have aggregated the BRFSS data over 7 years (1992–1998) and then smoothed the resulting maps to remove background noise and to highlight predominant within-state patterns in the data. These maps should prove more useful to state and local health officials and public health researchers than either state maps of very broad patterns or maps of the highly variable observed county data.

From the Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland,

Address correspondence and reprint requests to: Linda Williams Pickle, PhD, Statistical Research and Applications Branch, Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, 6116 Executive Blvd., Suite 504, Bethesda, MD 20892-8317. E-mail: PICKLEL@mail.nih.gov.

The full text of this article is available via AJPM Online at www.elsevier.com/locate/ajpmonline.

Methods

The BRFSS is a telephone-based nationwide survey conducted by the states, coordinated by the Centers for Disease Control and Prevention (CDC). The program began in 1984, and by

0749-3797/02/$–see front matter
PII S0749-3797(01)00402-0

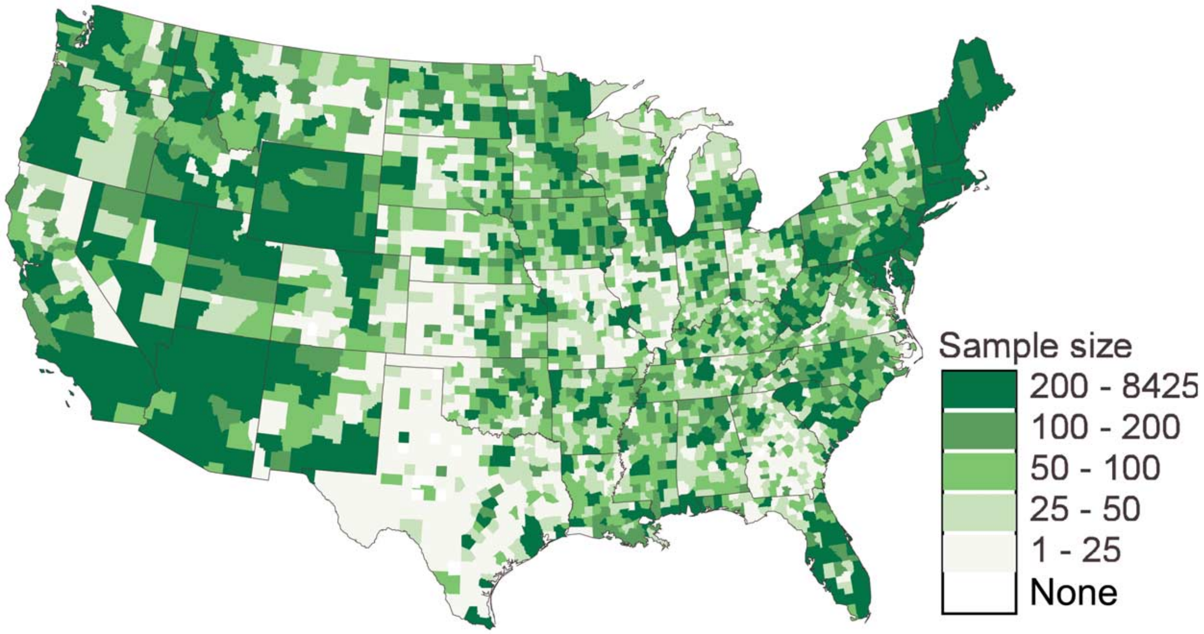


Figure 1. Sample sizes for Behavioral Risk Factor Surveillance System (BRFSS), combined data for 1992–1998.

1994 all U.S. states and the District of Columbia (hereafter referred to as "states") participated.[13–15] Each state uniformly administers a core questionnaire and may add one or more standardized supplements plus their own questions. The core and optional components include questions about the behaviors and conditions that place adults at risk for commonly occurring chronic diseases, injuries, and preventable infectious diseases. The targeted population is each state's noninstitutionalized civilian adults who live in a household with a telephone. The median response rate of the participating states, as defined by the Council of American Survey Research Organizations (CASRO), ranged from 70% in 1994 to 59% in 1998.[16] Surveys are conducted monthly using one of several survey methods, then data are aggregated to provide estimates representative of each state's population that year. Sample weights are provided that take into account the probability of sample selection and the possibly different sample designs. These weights plus uniform procedures for

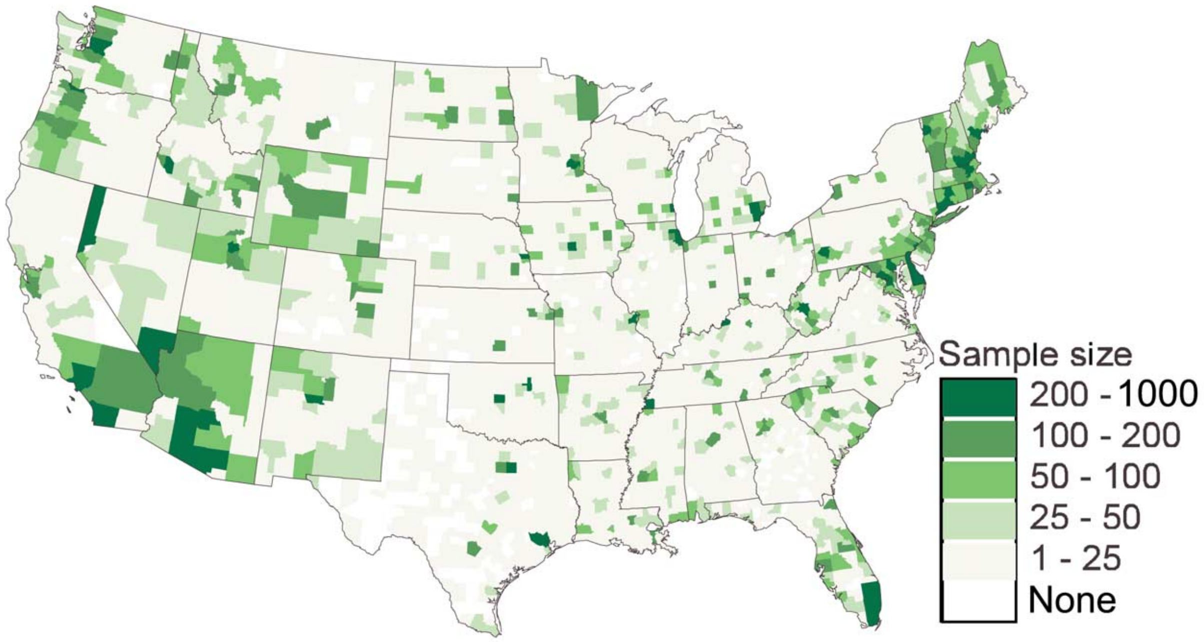


Figure 2. Sample sizes for females, aged 50–64 years, Behavioral Risk Factor Surveillance System (BRFSS), combined data for 1992–1998.

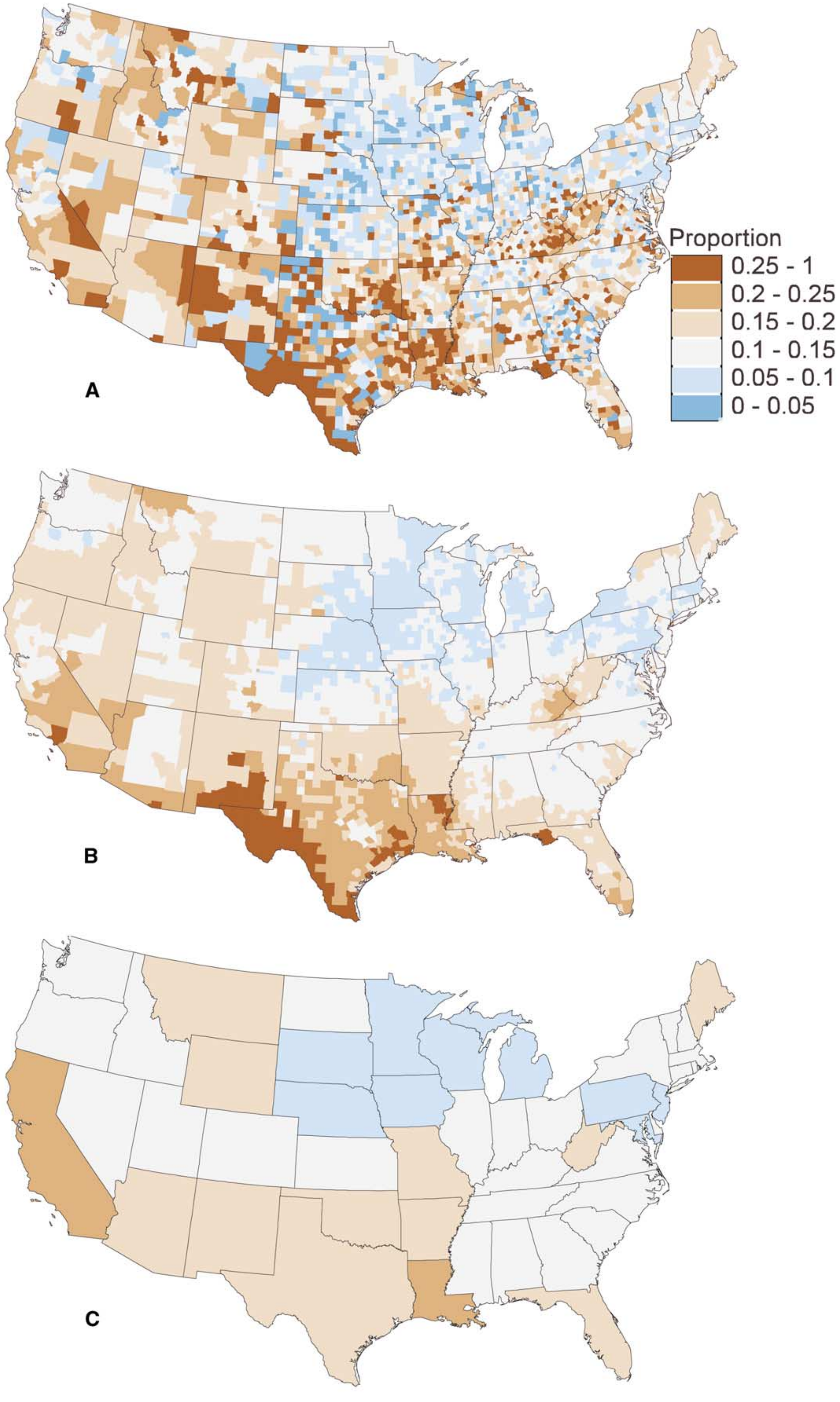


Figure 3. Weighted mean proportions of residents without healthcare coverage, Behavioral Risk Factor Surveillance System (BRFSS) 1992–1998 by county compared to BRFSS 1995 by state: (A) raw data, (B) smoothed data, (C) state data.

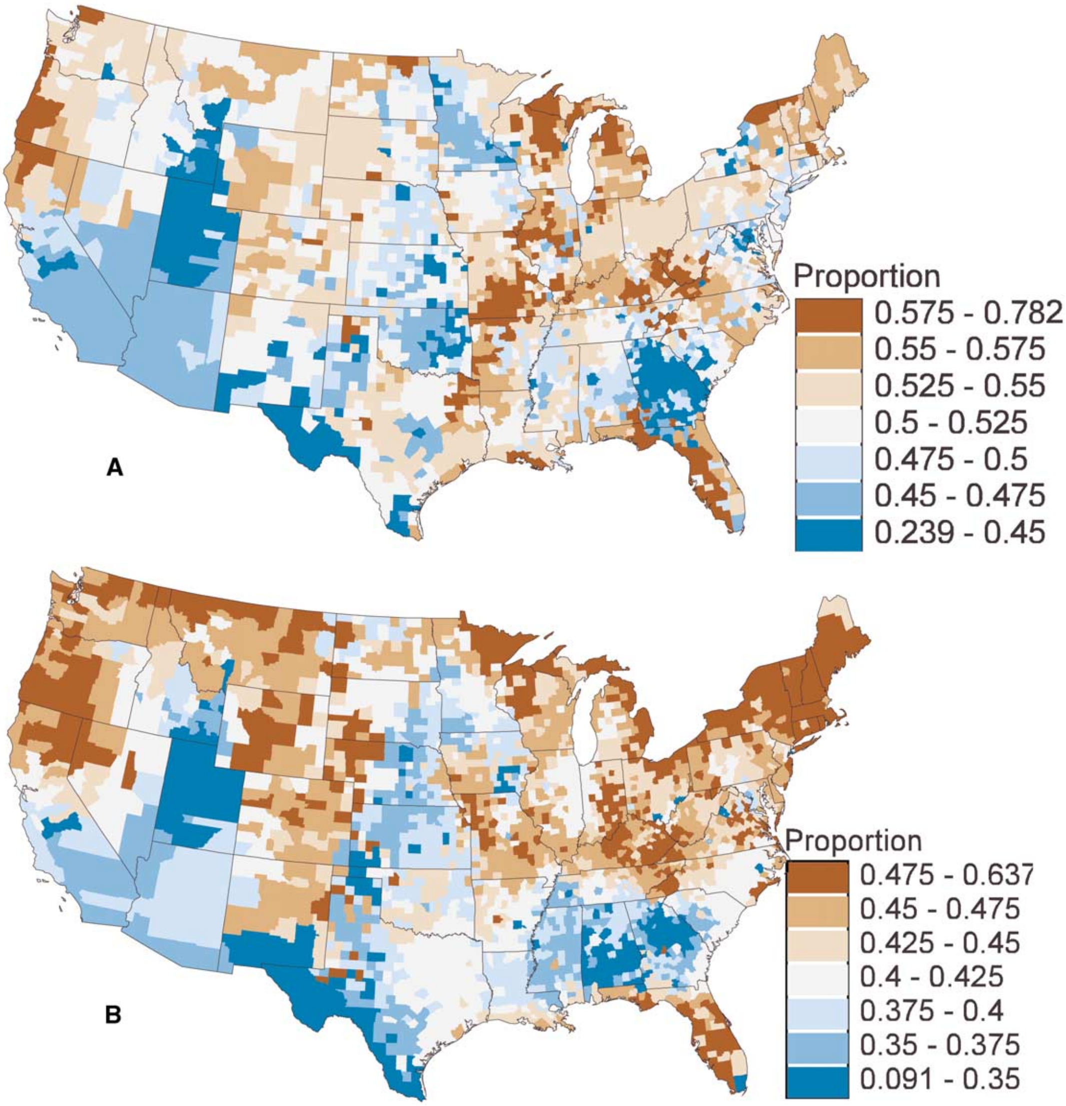


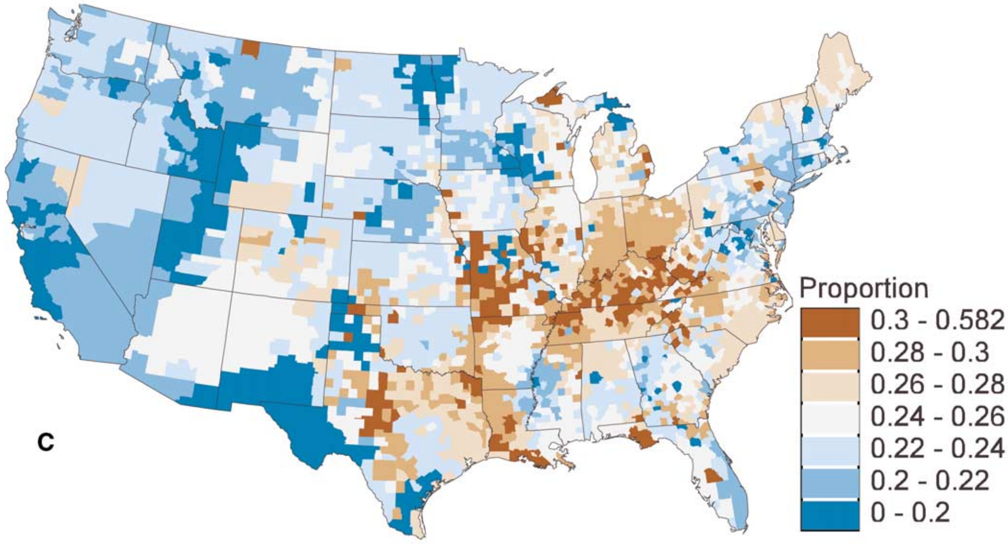
Figure 4. Smoothed maps of proportions of men and women who ever or currently smoke cigarettes: (A) proportion of men who ever smoked cigarettes, (B) proportion of women who ever smoked cigarettes, (C) proportion of men who currently smoke cigarettes, (D) proportion of women who currently smoke cigarettes.

data collection and processing provide the basis for data comparability across states. More detail on the survey design may be found at www.cdc.gov/nccdphp/brfss/usersguide.htm.

We aggregated BRFSS data for 1992–1998 by county for selected questions that had been asked consistently and uniformly over that time. Specifically, we computed proportions of respondents who reported that they currently or had ever smoked cigarettes; were at risk of obesity (body mass index [BMI] >120% of median reported BMI); had no healthcare plan; and, for females aged 50–79 years, had a mammogram during the past 2 years. Mammography patterns were examined separately for ages 50–64 and 65–79 years so that differences owing to Medicare coverage could be assessed. The relative definition of obesity allowed the absolute BMI cutoff to vary as the median BMI increased over our time period. Sample design weights were used to calculate these proportions to provide estimates representative of each state's targeted population. The data were pooled over time to provide more reliable estimates by county, although the aggregated data thus cannot be used to examine changing patterns over time.

There were 823,489 original respondents to the BRFSS during 1992–1998. Respondents from three states or territories were deleted: Alaska (n=11,153) because it had no county designation; Puerto Rico (n=6426) because it was not sampled for over half of the study time period; and Hawaii (n=14,828) because over half of its respondents were from a single county (Honolulu). With no neighboring counties to supplement their own results, no within-state patterns could be determined for these states. In addition to these deletions, 86,129 respondents did not name their county and the county designation was masked on 38,223 records for counties with fewer than 50 respondents in a single year, leaving 666,730 respondents for analysis.

Sample sizes ranged from 8459 to 27,913 by state, with a median of 14,280. The median sample size per county of the



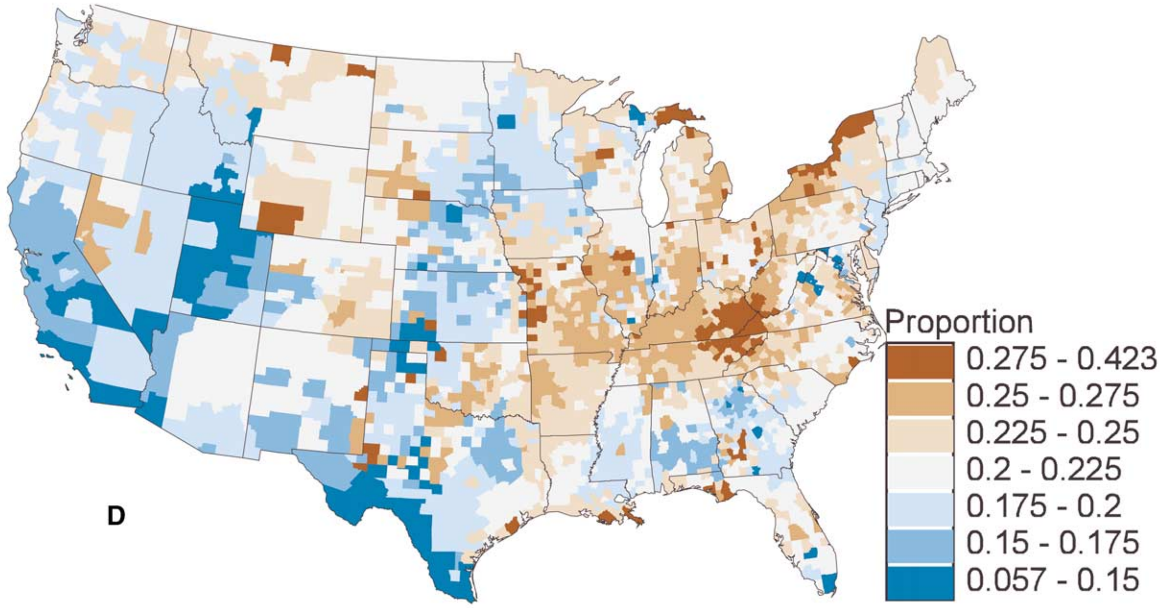


Figure 4. *(continued)*

coterminous states was 60, which represented 1514 state residents (i.e., the weighted sample size). As shown in Figure 1, the sample sizes were roughly proportional to the population, with small samples taken in sparsely populated counties. Samples for males and females were of approximately equal size and followed the geographic patterns shown in Figure 1 (not shown). Sample sizes for females aged 50–64 (and 65–79) years were naturally smaller than those for all residents (Figure 2).

Because of the relatively small samples in many counties, patterns in maps of the original data appeared quite scattered. In order to reveal underlying geographic patterns in the data, we smoothed the proportions by a weighted two-dimensional median-based smoothing algorithm, termed "headbanging," implemented as follows.[17] For each county, up to 30 neighboring counties comprised its "smoothing window," which was of sufficient size to begin to show regional patterns while retaining patterns apparent in the raw data maps. The medians of the higher and lower 50% of the neighbors were calculated (the "high median" and "low median"), weighted by county population. If the observed center county value was between the high and low medians, or its population was much greater than its neighbors, its value was not changed. Otherwise, if the center value was less than (greater than) the low (high) median, its value was changed to equal the low (high) median. This process was repeated 10 times across the map. This process can be thought of as a roughly circular moving average (median) of neighboring counties applied to each county in turn across the country.

Geographic smoothing algorithms "borrow information" from neighboring areas to stabilize results from sparsely populated areas. This reduces the variability in the data, allowing patterns to emerge, but increases the bias in the estimates for each small area. Consequently, the reader should not attempt to interpret the proportions for any single county. The variance reduction, however, allows the reader to identify and compare clusters of counties with similar values. To illustrate, the mean proportion of current male smokers within a band of counties along the Mexican border is 30%, compared to 20% within an adjacent band of counties. The sample sizes of these county clusters, 1468 and 3218, respectively, would generate clearly non-overlapping 95% confidence limits. Although these maps are presented for exploratory, not inferential, purposes, this example illustrates the strength gained from smoothing.

Use of population weights in the smoothing process ensured that unusually high or low proportions that were reliable owing to large populations were not modified, whereas values based on sparse populations were modified to be more like those of the surrounding counties. For the 62 counties with no sampled residents during 1992–1998, we arbitrarily substituted the state weighted mean proportion, but these values were then smoothed to be similar to their more populous neighboring counties.

The resulting smoothed proportions are shown as categorized area-shaded maps.[18] Cut points for the lowest and highest classes were approximately the 5th and 95th percentiles of each factor's distribution, respectively. Intervening classes were of equal length—5% each for health insurance and mammography, 2% for male current smoking, and 2.5% each for obesity and the other smoking maps—resulting in six or seven categories for each map, depending on the range of the proportions. For example, cut points were 5%, 10%, 15%, 20%, and 25% for no health insurance (Figure 3). These categories were color-coded according to whether the level of the risk factor was worse (red-brown) or better (blue) than the mean (gray) of the smoothed county values. For example, the darkest red-brown indicates the worst level of the risk factor, i.e., more smokers and obese residents and fewer residents who were covered by health plans or who had recent mammograms.

Results

A map of the original proportions of residents with no healthcare coverage looked quite scattered (Figure 3a). After smoothing (Figure 3b), it can be seen that fewer residents have health insurance along the Texas–Mexico border and in portions of Louisiana, Oklahoma, and Kentucky. In general, health plan coverage is more common in the Northeastern and Midwestern states than in the South and West. This broad pattern can be seen in the published state map for 1995 (Figure 3c), but the smoothed county map shows heterogeneity within states (e.g., Texas and Kentucky).[19] The effects of smoothing and aggregation are apparent in the ranges of the proportions: 0% to 100% for the original data, 5% to 43% for smoothed county data, and 7.8% to 21% for state data (Figure 3).

Because different patterns for lung cancer mortality rates have been identified elsewhere for white men and women,[20] the smoking maps were stratified by gender. As seen in Figure 4, after smoothing, the proportions of those who currently or had ever smoked were relatively high in eastern Kentucky, southern and western Missouri, and portions of several Rocky Mountain states for both men and women. Counties along the northern Pacific Coast and in upstate New York had a higher proportion of residents who had ever smoked, particularly women. The smoking patterns by county were more strongly correlated with age-adjusted lung cancer mortality rates for 1995–1996 after smoothing than before ($r = 0.29$ vs 0.11), consistent with our goal of enhancing relevant risk factor patterns.

Figure 5 shows smoothed maps for the proportion of women, aged 50–64 and 65–79 years, who reported having a mammogram during the last 2 years. For both age groups, fewer than half of the women had a mammogram in western New York, Pennsylvania, Mississippi, and parts of West Virginia, Kentucky, Tennessee, Florida, Texas, Kansas, Nebraska, Montana, and Idaho. Notably high proportions of women were screened in New England, Michigan, Alabama, and Wyoming. Some differences by age group are noted in the Midwest and along the Pacific coast.

Obesity was more prevalent in many of these same areas, with particularly high rates in West Virginia, eastern Kentucky, Tidewater Virginia, eastern Michigan, southern Alabama and Georgia, and southern Texas (Figure 6).

Discussion

The purpose of the maps presented here is to highlight within-state geographic patterns of selected health behaviors related to a number of chronic diseases. For nearly a decade, the BRFSS has provided the only source of risk factor data collected in a systematic way for counties across the United States. These data are useful for ecologic studies, for identifying high-risk communities where health promotion or disease-prevention programs could be implemented or evaluated, and for educating the public, healthcare professionals, and legislators.[13,14] Similar information may be obtained from personal interviews of individuals, but time and cost constraints limit these studies to very localized areas. Information about health insurance coverage or screening procedures is available in some administrative databases, such as hospital records, but it is often incomplete and not representative of the resident population. On a broader scale, several surveys provide estimates of risk factor prevalence, but only for the nation, for large geographic regions, or for large population areas (e.g., National Health Interview Survey, National Health and Nutrition Examination Survey, Current Population Survey [CPS]).

Patterns identified by smoothing the BRFSS data are consistent with published state-level maps of these data and data from other surveys. For example, the highest proportion of residents without health insurance has

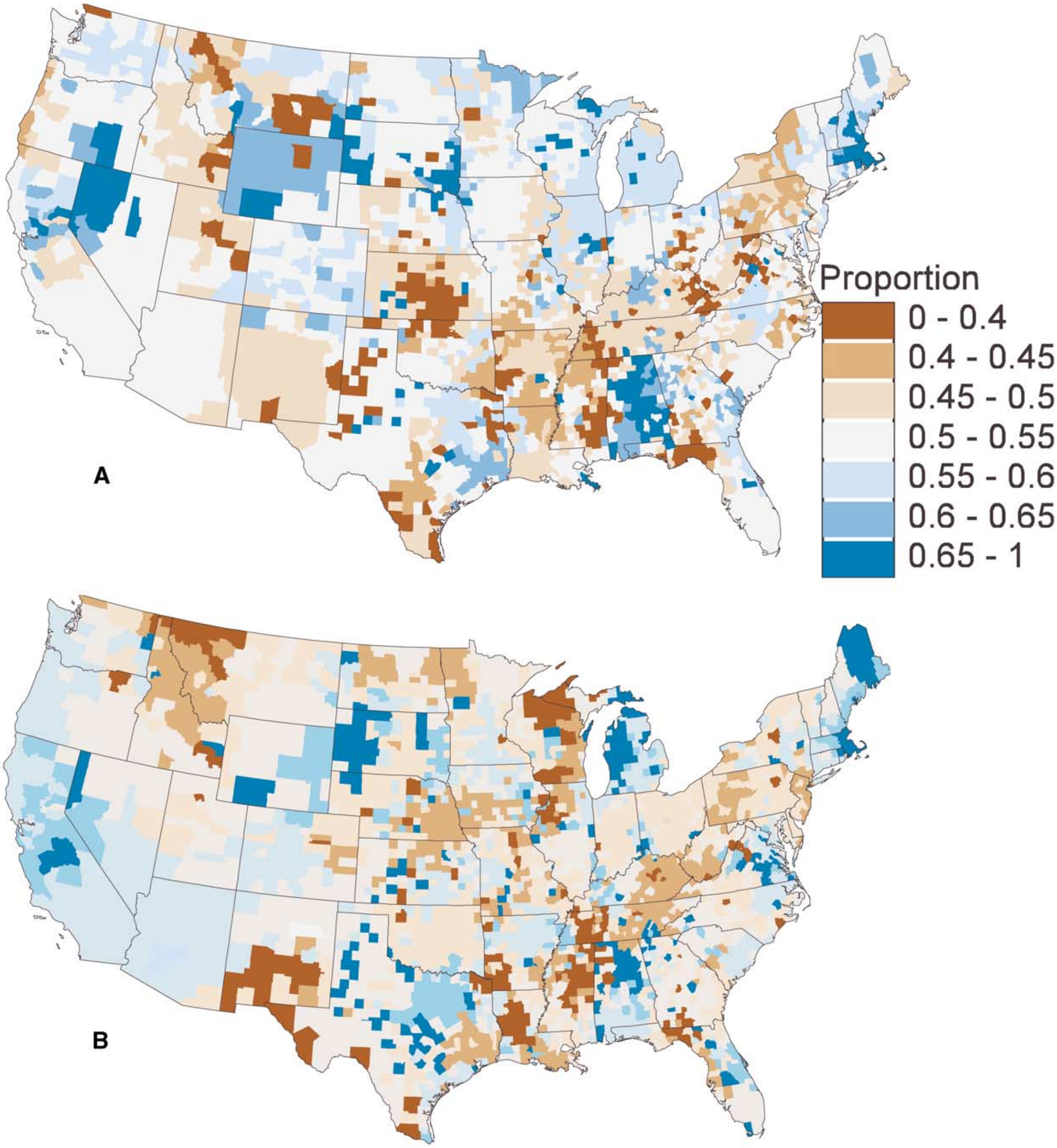


Figure 5. Smoothed maps of the proportion of women, aged 50–64 and 65–79 years, who had a mammogram in the past 2 years: (A) ages 50–64, (B) ages 65–79.

been reported in a band of southern states from Mississippi to California/Nevada.[5,21] The smoothed county maps shown here indicate that only portions of these states have low coverage and, furthermore, show a cluster of counties with low coverage in eastern Kentucky/western West Virginia, a pattern that is masked at the state level.

The maps of obesity and current smoking among males resemble a map of low income, with relatively high levels along portions of the Mississippi River and into Kentucky (see p. 208 in Reference 20), similar to BRFSS state maps of both risk factors.[8] All four smoking maps show a relatively high level in West Virginia, Kentucky, eastern Tennessee, and Missouri. Many residents of Florida and the northern Pacific coast, both male and female, were at one time cigarette smokers, but these areas now have a lower proportion of current smokers. These maps appear similar to maps of lung cancer mortality, particularly for women,[20] and to previously published state maps of smoking data from the CPS.[22] However, one comparison of state smoking data from the BRFSS and CPS surveys showed several

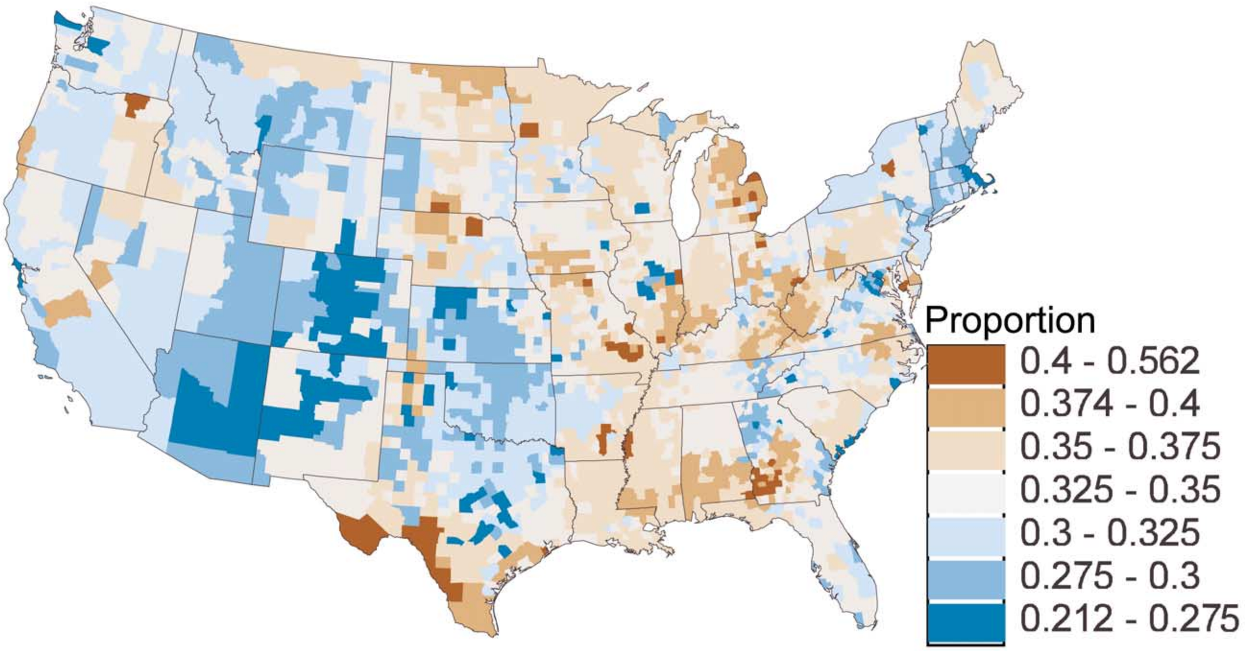


Figure 6. Smoothed map of the proportion of residents who were at risk of obesity (>120% of median body mass index).

states where current smoking prevalence was significantly different as measured by the two surveys.[23]

An ecologic analysis of national survey data has shown mammography use to be positively correlated with income and negatively correlated with the proportion of the population of Hispanic origin.[24] Our mammography patterns appear to be positively correlated with income, but show relatively high screening rates along the Mexican border, an area with many Hispanic residents. The map for older women closely resembles a map of mammography use based on Medicare records.[10,25] Differences between the maps for younger and older women are likely owing to coverage of preventive services by private insurance compared to Medicare.

Overall, estimates from the BRFSS agree well with those from in-person interviews and physical measurements.[26,27] However, the reader must use caution in interpreting the smoothed within-state patterns. The sample sizes for most counties are small, although the consistencies with other work suggests that realistic patterns have emerged for clusters of counties by smoothing the data. An alternative to smoothing the maps is a statistical model of the risk factor data, but this requires that relevant predictors be identified. Furthermore, a model-based approach would suffer more from the small sample sizes of many counties than the method used here unless spatial smoothing is included through correlation among adjacent counties.

The purpose of these maps is to identify within-state patterns of health risk factors that may interest public health researchers. However, smoothing can blur sharp differences that occur at artificial boundaries such as states. Thus, if the mapped pattern could be the result of governmental policies, such as free mammograms, then published state maps of observed proportions should be used rather than smoothed county maps.

Potential sources of bias in BRFSS data include self-reporting, telephone coverage, and survey nonresponse. Self-reports of cigarette smoking prevalence have been found to be within 1%–2% of results from physiological measurements but obesity is typically underreported.[26–30] Comparison of self-reports of mammograms to radiology center logs showed that women often report that their examination was more recent than in actuality, which would result in an overestimate of the proportion screened within a specified period, as mapped here.[31] However, there is no indication that these biases vary geographically.

Surveys such as BRFSS are often criticized because of the potential bias due to excluding persons with no telephone. Overall, only 5% of U.S. residents have no telephone, but the proportion without telephones is greater for blacks (10%) and for those living below the poverty level (17%).[32] However, comparison of health behaviors between those with and without telephones shows little difference; the absolute error for the factors mapped here was at most 1.2%, even when comparisons were restricted to those living below the poverty level.[32] Coverage for some subgroups, such as American Indians living on reservations in the southwest, may be much worse.[33] Thus, the maps presented here are applicable for general comparisons but may not be representative for a specific subgroup that has very low telephone coverage.

Of greater concern is a potential bias due to survey nonresponse, which has increased over time. A recent study demonstrated that response rates could be increased substantially by rigorous follow-up, which altered estimates of income, health insurance coverage, and race by as much as 4%.[34]

Despite these limitations, the resulting maps are consistent with other reports. The smoothing algorithm has identified within-state patterns of health risk factors not apparent in previously published state maps and, thus, may prove useful in understanding localized patterns of health risk behaviors and in targeting intervention activities in the United States.

The authors would like to thank Dr. David E. Nelson for background information and insightful comments on an earlier draft of this manuscript.

Electronic files for the figures may be downloaded from srab. cancer.gov/publications/GIS.

References

1. Surgeon General of the United States. The health consequences of smoking. A public health service review. Washington, DC: U.S. Government Printing Office, 1967.
2. Troiano RP, Frongillo EA Jr, Sobal J, Levitsky DA. The relationship between body weight and mortality: a quantitative analysis combining information from existing studies. Int J Obes 1996;20:63–75.
3. Foster WR, Burton BT, eds. National Institutes of Health Consensus Development Panel on the health implications of obesity. Ann Intern Med 1985;103(Suppl 6):1073–7.
4. Schottenfeld D, Fraumeni JF Jr. Cancer epidemiology and prevention, 2nd edition. New York: Oxford University Press, 1996:451.
5. Salem N. Health insurance coverage and receipt of preventive health services—United States, 1993. MMWR Morb Mortal Wkly Rep 1995;44:219–25.
6. Weissman JS, Stern R, Fielding SL, Epstein AM. Delayed access to health care: risk factors, reasons, and consequences. Ann Intern Med 1991;114:325–31.
7. Himmelstein DU, Woolhandler S. Care denied: U.S. residents who are unable to obtain needed medical services. Am J Public Health 1995;85:341–4.
8. Hahn RA, Heath GW, Chang MH. Cardiovascular disease risk factors and preventive practices among adults—United States, 1994: A behavioral risk factor atlas. MMWR CDC Surveill Summ 1998;47:35–69.
9. Ashton CM, Petersen NJ, Souchek J, et al. Geographic variations in utilization rates in Veterans Affairs hospitals and clinics. N Eng J Med 1999;340:32–9.
10. Wennberg JE, Cooper MM, eds. The Dartmouth atlas of health care in the United States. Chicago, IL: American Hospital Publishing, Inc., 1996.
11. Welch PE, Miller ME, Welch HG, Fisher ES, Wennberg JE. Geographic variation in expenditures for physicians' services in the United States. N Engl J Med 1993;328:621–7.
12. Diez-Roux AV, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. N Engl J Med 2001;345:99–106.
13. Nelson DE, Holtzman D, Waller M, Leutzinger CL, Condon K. Objectives and design of the Behavioral Risk Factor Surveillance System. 1998 Proceedings of the Section on Survey Research Methods, American Statistical Association. Alexandria, VA: American Statistical Association, 1998:214–8.
14. Powell-Griner E. Uses and limitations of the Behavioral Risk Factor Surveillance System data. 1998 Proceedings of the Section on Survey Research Methods, American Statistical Association. Alexandria, VA: American Statistical Association, 1998:219–23.
15. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System (BRFSS). Available at: www.cdc.gov/nccdphp/brfss/. Accessed on August 13, 2001.
16. Centers for Disease Control and Prevention. 1998 BRFSS Summary Quality Control Report. Available at: www.cdc.gov/nccdphp/brfss/pdf/98quality.pdf. Accessed on August 13, 2001.
17. Mungiole M, Pickle LW, Simonson KH. Application of a weighted headbanging algorithm to mortality data maps. Stat Med 1999;18:3201–9.
18. Environmental Systems Research Institute, Inc. ArcView GIS version 3.2. Redlands, CA: Environmental Systems Research Institute, Inc., 1996.
19. BRFSS Coordinators. State-specific prevalence estimates of uninsured and underinsured persons—Behavioral Risk Factor Surveillance System, 1995. MMWR Morb Mortal Wkly Rep 1998;47:51–5.
20. Pickle LW, Mungiole M, Jones GK, White AA. Atlas of United States Mortality. DHHS Pub. No. (PHS) 97-1015. Hyattsville, MD: National Center for Health Statistics, 1996.
21. Bureau of the Census. Health insurance coverage: Who goes without health insurance? Publication P60-195. Washington, DC: Bureau of the Census, 1996.
22. Shopland DR, Hartman AM, Gibson JT, Mueller MD, Kessler LG, Lynn WR. Cigarette smoking among U.S. adults by state and region: estimates from the Current Population Survey. J Natl Cancer Inst 1996;88:1748–58.
23. Arday DR, Tomar SL, Nelson DE, Merritt RK, Schooley MW, Mowery P. State smoking prevalence estimates: a comparison of the Behavioral Risk Factor Surveillance System and Current Population Surveys. Am J Public Health 1997;87:1665–9.
24. Wells BL, Horm JW. Targeting the underserved for breast and cervical cancer screening: the utility of ecological analysis using the National Health Interview Survey. Am J Public Health 1998;88:1484–9.
25. Goodman DC, Wennberg JE. Maps and health: the challenges of interpretation. J Public Health Manage Pract 1999;5(4):xii–xvi.
26. Gentry EM, Kalsbeek WD, Hogelin GC, et al. The Behavioral Risk Factor Surveys: II. Design, methods, and estimates from combined state data. Am J Prev Med 1985;1:9–14.
27. Jackson C, Jatulis DE, Fortmann SP. The Behavioral Risk Factor Survey and the Stanford Five-city Project Survey: a comparison of cardiovascular risk behavior estimates. Am J Public Health 1992;82:412–6.
28. Pierce JR, Aldrich PN, Hanratty S, Dwyer T, et al. Uptake and quitting smoking trends in Australia 1974–1984. Prev Med 1987;16:252–60.
29. Fortmann SP, Rogers T, Vranizan K, Haskell WL, et al. Indirect measures of cigarette use: Expired-air carbon monoxide versus plasma thiocyanate. Prev Med 1984;13:127–35.
30. Bowlin SJ, Morrill BD, Nafziger AN, Lewis C, Pearson TA. Reliability and changes in validity of self-reported cardiovascular disease risk factors using dual response: the Behavioral Risk Factor Survey. J Clin Epidemiol 1996;49:511–7.
31. Degnan D, Harris R, Ranney J, Quade D, Earp JA, Gonzalez J. Measuring the use of mammography: Two methods compared. Am J Public Health 1992;82:1386–8.
32. Anderson JE, Nelson DE, Wilson RW. Telephone coverage and measurement of health risk indicators: data from the National Health Interview Survey. Am J Public Health 1998 Sep;88:1392–5.
33. Pearson D, Cheadle A, Wagner E, Tonsberg R, Psaty BM. Differences in sociodemographic, health status, and lifestyle characteristics among American Indians by telephone coverage. Prev Med 1994;23:461–4.
34. Keeter S, Miller C, Kohut A, Groves RM, Presser S. Consequences of reducing nonresponse in a national telephone survey. Public Opinion Q 64:125–48.