

## Cigarette Use and Race/Ethnicity Dataset Introduction

### **Abstract**

This is a subset of the data from two Supplements to the Current Population Survey conducted by the U.S. Census Bureau. Soulakova and Crockett (2023) investigated the joint role of race/ethnicity and health insurance coverage and their relationship to current smoking among adults in the United States. The data in this subset ( $n=39,882$ ) represent six racial/ethnic groups, in addition to several other factors including demographic and socioeconomic characteristics such as age, biological sex, highest level of education, disability status, and health insurance coverage (Private, Medicare, and Medicaid). Because the source data were collected using a multi-stage sampling design, it is recommended that statistical analyses utilize the survey weights including the main weight and 160 replicate weights included in the dataset (total weight,  $N=218,443,483$ ). The data are clean and there are no missing values.

### **Background**

It is well known that substance use, in particular smoking, leads to adverse health outcomes and poses many public health challenges. With respect to smoking, type of health insurance coverage is an established correlate, making the clinical encounter a potential opportunity for smoking cessation counseling. Less well understood, however, is the joint role of race/ethnicity and health insurance coverage in smoking behaviors. If the association of health insurance coverage and current smoking is different for distinct race/ethnicity populations (effect modification), the implication is that, within the same health insurance plan, some racial/ethnic groups may be benefitting less from counseling and require additional interventions.

### **Study Objective**

The goal of this study was to identify variations among six (6) racial/ethnic populations in the association between type of health insurance coverage (Private, Medicare, Medicaid) and smoking (defined as the binary outcome “current smoking status” among U.S. adults 18+ years old), overall and after adjustment for selected covariates (e.g., biological sex, age, employment, marital status, and disability status). Soulakova and Crockett (2023) hypothesized that even after adjusting for other factors that may influence smoking behaviors among adults, the relationship between health insurance coverage type and current cigarette use is different depending on race/ethnicity.

### **Study Design**

This dataset is a subset of a cross-sectional complex survey. This subset was created by merging data from two Supplements of the Current Population Survey: the Tobacco Use Supplement (TUS, January and May 2019, US Department of Commerce, Census Bureau, 2019) and the 2019 Annual Social and Economic Supplement (ASEC, Bureau of the Census for the Bureau of Labor Statistics, 2019). To merge the TUS and ASEC data, Soulakova and Crockett (2023) used the linkage procedure described elsewhere (Reyes-Guzman, Patel, Wang, et al., 2023). The merged dataset includes responses from 39,882 adults who self-identified as Hispanic, non-Hispanic (NH) American Indian/Alaska Native, NH Asian/Asian American, NH Black/African American, NH Hawaiian/Pacific Islander, or NH White. Respondents who reported more than one race (i.e., multiracial) were not included in the study. In addition, only respondents for whom complete information on all study measures was available were included in the merged data. Thus, the dataset includes no missing values.

### **Subjects & Variables**

n = 39, 882 observations

p = 13 study variables plus 161 survey weight variables, including one main weight and 160 replicate weight variables

The 13 study variables include current use of combustible cigarettes (SMOKSTAT), 3 indicator variables describing health insurance coverage (NOW\_MCAID, NOW\_MCARE, NOW\_PRIV) plus 9 variables measuring demographics and socioeconomic status.

### **Additional Information**

Statistical analysis was conducted using Survey Package in SAS®9.4 (SAS Institute Inc. 2013). Some key principles of analyses of the TUS data are outlined in Ha and Soulakova (2018). For your reference, the analyses of this dataset utilize values of the Fay coefficient specified by Reyes-Guzman et al. (2023), which is different from the one used for analyses of TUS data (Ha and Soulakova, 2018).

### **Citations**

Soulakova, J.N. and Crockett, L.J., 2023, November. *Current Use of Cigarettes in the United States: The Joint Role of Race/Ethnicity and Health Insurance Coverage*. In *Healthcare* (Vol. 11, No. 23, p. 3014). MDPI. Available at <https://www.mdpi.com/2227-9032/11/23/3014>

Reyes-Guzman CM, Patel M, Wang TW, Corcy N, Chomenko D, Slotman B, et al. *Disparities in Smokefree and Vapefree Home Rules and Smokefree Policy Attitudes Based on Housing Type and Cigarette Smoking Status, United States*, 2019. *Int. J. Environ. Res. Public Health* [Internet]. NLM (Medline); 2023; 20:6356. Available at <https://www.mdpi.com/1660-4601/20/14/6356/htm>

Ha T, Soulakova JN. Importance of Adjusting for Multi-Stage Design when Analyzing Data from Complex Surveys. In: Zhao Y, Chen D-G, editors. *New Front. Biostat. Bioinformatics*. Springer Nature Switzerland; 2018. p. 257–68. Available at [https://link.springer.com/chapter/10.1007/978-3-319-99389-8\\_12](https://link.springer.com/chapter/10.1007/978-3-319-99389-8_12)

Bureau of the Census for the Bureau of Labor Statistics. 2019 Annual Social and Economic (ASEC) Supplement. Washington: U.S. Census Bureau, 2019.” Available at <https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar19.pdf>

“US Department of Commerce, Census Bureau (2019). National Cancer Institute and Food and Drug Administration Co-Sponsored Tobacco Use Supplement to the Current Population Survey.” 2019. Available at <https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/>

SAS Institute Inc. SAS® 9.4 Product Documentation. Cary, NC, USA SAS Inst. Inc. 2013.