

---

# Behavioral Risk Factor Surveillance System

OVERVIEW: BRFSS 2013

August 15, 2014





## Background

The Behavioral Risk Factor Surveillance System (BRFSS) is a collaborative project between all of the states in the United States (US) and participating US territories and the Centers for Disease Control and Prevention (CDC). The BRFSS is administered and supported by CDC's Population Health Surveillance Branch, under the Division of Population Health at the National Center for Chronic Disease Prevention and Health Promotion. BRFSS is an ongoing surveillance system designed to measure behavioral risk factors for the non-institutionalized adult population (18 years of age and older) residing in the US. The BRFSS was initiated in 1984, with 15 states collecting surveillance data on risk behaviors through monthly telephone interviews. Over time, the number of states participating in the survey increased; by 2001, 50 states, the District of Columbia, Puerto Rico, Guam, and the US Virgin Islands were participating in the BRFSS. Today, all 50 states, the District of Columbia, Puerto Rico, and Guam collect data annually and American Samoa, Federated States of Micronesia, and Palau collect survey data over a limited point- in-time (usually one to three months). In this document, the term “state” is used to refer to all areas participating in BRFSS, including the District of Columbia, Guam, and the Commonwealth of Puerto Rico.

**The BRFSS objective is to collect uniform, state-specific data on preventive health practices and risk behaviors that are linked to chronic diseases, injuries, and preventable infectious diseases that affect the adult population.** Factors assessed by the BRFSS in 2013 include tobacco use, HIV/AIDS knowledge and prevention, exercise, immunization, health status, healthy days — health-related quality of life, health care access, inadequate sleep, hypertension awareness, cholesterol awareness, chronic health conditions, alcohol consumption, fruits and vegetables consumption, arthritis burden, and seatbelt use. Since 2011, BRFSS conducts both landline telephone- and cellular telephone-based surveys. In conducting the BRFSS landline telephone survey, interviewers collect data from a randomly selected adult in a household. In conducting the cellular telephone version of the BRFSS questionnaire, interviewers collect data from an adult who participates by using a cellular telephone and resides in a private residence or college housing.

BRFSS field operations are managed by state health departments that follow protocols adopted by the states with technical assistance provided by CDC. State health departments collaborate during survey development, and conduct the interviews themselves or by using contractors. The data are transmitted to the CDC for editing, processing, weighting, and analysis. An edited and weighted data file is provided to each participating health department for each year of data collection, and summary reports of state-specific data are prepared by the CDC. Health departments use the data for a variety of purposes, including identifying demographic variations in

health-related behaviors, targeting services, addressing emergent and critical health issues, proposing legislation for health initiatives, and measuring progress toward state health objectives.<sup>1</sup> For specific examples of how state officials use the finalized BRFSS data sets, please refer to the appropriate [state information](#) on the BRFSS Web site.

Health characteristics estimated from the BRFSS pertain to the non-institutionalized adult population, aged 18 years or older, who reside in the US. In 2013, additional question sets were included as optional modules to provide a measure for several childhood health and wellness indicators, including asthma prevalence for people aged 17 years or younger.

As noted above, respondents are identified through telephone-based methods. Overall, an estimated 97.5% of US households had telephone service in 2012.<sup>2</sup> Telephone coverage varies across states with a range of 95.3% in New Mexico to 98.6% in Connecticut. The increasing percentage of households that are abandoning their landline telephones for cellular telephones has significantly eroded the population coverage provided by landline telephone-based surveys to pre-1970s levels. For the first half of 2013, the percentage of cellular telephone-only households was 39.4 percent.<sup>3</sup> This is an increase of 1.2 percentage points over the preceding 6-month period. In households where both landline telephone and wireless telephone service is available, there is a trend toward increased use of wireless communication. In 2013, BRFSS respondents who received 90 percent or more of their calls on cellular telephones were eligible for participation in the cellular telephone survey.

No direct method of accounting for non-telephone coverage is employed by the BRFSS. Continuing a weighting method first introduced in 2011, BRFSS used the weighting methodology called iterative proportional fitting (or “raking”) to weight the data. Raking adjusts the data so that groups underrepresented in the sample can be more accurately represented in the final data set. Raking allows for the incorporation of cellular telephone survey data; it permits the introduction of additional demographic characteristics and more-accurately matches sample distributions to known demographic characteristics of populations, as compared with the pre-2011 BRFSS weighting methods. The use of raking has been shown by researchers to reduce error within estimates<sup>4</sup>. BRFSS raking includes categories of age by gender, detailed race and ethnicity groups, education levels, marital status, regions within states, gender by race and ethnicity, telephone source, renter/owner status, and age groups by race and ethnicity. In 2013, 50 states, the District of Columbia, Guam, and Puerto Rico collected samples of interviews conducted both by landline telephone and cellular telephone.

## BRFSS Design

### The BRFSS Questionnaire

Each year, the states represented by their BRFSS coordinators and CDC agree on the content of the survey questionnaire. The BRFSS questionnaire consists of a core component and optional modules. Many questions are taken from established national surveys, such as the National Health Interview Survey or the National Health and Nutrition Examination Survey. This practice allows BRFSS to take advantage of questions that have been tested and allows states to compare their data with those from other surveys. Any new questions that states, federal agencies or other entities propose as additions to BRFSS must go through cognitive testing and field testing before they can become part of the BRFSS questionnaire. In addition, a majority vote of all state representatives is required before questions are adopted. BRFSS guidelines, agreed upon by the state representatives and CDC, specify that all states ask the core component questions without modification; they may choose to add any, all, or none of the optional modules and may add questions of their choosing at the end of the questionnaire (state-added questions).

The questionnaire has three parts:

**i. Core component:** A standard set of questions that all states use. Core content includes queries about current health-related perceptions, conditions, and behaviors (e.g., health status, health care access, alcohol consumption, tobacco use, disability, and HIV/AIDS risks), as well as demographic questions. The core component includes the annual core comprised of questions asked each year, and rotating core questions which are included in even- and odd –numbered years.

**ii. Optional BRFSS modules:** Sets of questions on specific topics (such as excess sun exposure, cancer survivorship, mental illness, and stigma) that states elect to use on their questionnaires. In 2013, 22 optional modules were supported by the BRFSS. Generally, CDC programs submit module questions and the states vote to adopt final questions that can be included as optional modules. For more information, please see the [questionnaire section](#) of the BRFSS Web site.

**iii. State-added questions:** Individual states develop or acquire these questions and add them to their BRFSS questionnaires. CDC neither edits nor evaluates these questions.

BRFSS supported 22 modules in 2013, but states limited modules and added questions to only the most useful for their state program purposes, in order to keep surveys at a reasonable length. Because different states have different needs, there is wide variation between states in terms of question totals each year. BRFSS implements

a new questionnaire in January and usually does not change it significantly for the rest of the year. The flexibility of state-added questions, however, does permit additions, changes, and deletions at any time during the year.

The 2013 list of [optional modules](#) used on both the landline telephone and cellular telephone surveys is available on the BRFSS Web site. In order to allow for a wider range of questions in optional modules, combined landline telephone and cellular telephone data for 2013 include up to 3 split versions of the questionnaire. A split version is when a subset of telephone numbers for data collection still followed the state sample design, and administrators used it as the state's BRFSS sample, but the optional modules and state-added questions may have been different from other split version questionnaires. For additional information on split version questionnaires see the document [2013 Combined Landline Telephone and Cellular Telephone Survey Multiple-Version Questionnaire Use of Data](#).

### **Annual Questionnaire Development**

The governance of the BRFSS includes a representative body of state health officials, elected by region. During the year, the State BRFSS Coordinators Working Group meets with CDC's BRFSS program management. One task of this group is to develop a 5-year, long-term plan for the BRFSS core instrument. The 2013 BRFSS questionnaire represents the third year of a 5-year plan (2011 – 2015) .

Before the beginning of the calendar year, the CDC provides states with the text of the core component and the optional modules that the BRFSS will support in the coming year. States select their optional modules and ready any state-added questions they plan to use. Each state then constructs its own questionnaire. The order of the questioning is always the same: interviewers ask questions from the core component first; then they ask any questions from the optional modules, and the state-added questions. This content order ensures comparability across states and follows BRFSS guidelines. Generally, the only changes that the standard protocol allows are limited insertions of state-added questions on topics related to core questions. The CDC and state partners must agree to these exceptions. However, in some cases, states have not been able to follow all set guidelines. Users should refer to the [Comparability of Data](#) document, which lists the known deviations.

Once each state finalizes its questionnaire content—consisting of the core questionnaire, optional modules, and state-added questions--the state prepares a hard-copy or electronic version of the instrument and sends it to the CDC. States use the [questionnaire](#) without changes for one calendar year, and CDC archives a copy on the BRFSS Web site. If a significant portion of any state's population does not speak English, states have the option of translating the questionnaire into other languages. Currently, the CDC provides a Spanish version of the core questionnaire and optional modules. Specific wording of the Spanish version of the questionnaire may be adapted by the states to fit the needs of their Hispanic populations.

## **Sample Description**

In a telephone survey such as the BRFSS, a sample record is one telephone number in the list of all telephone numbers the system randomly selects for dialing. To meet the BRFSS standard for the participating states' sample designs, one must be able to justify sample records as a probability sample of all households with telephones in the state. All participating areas met this criterion in 2013. Fifty-one projects used a disproportionate stratified sample (DSS) design for their landline samples. Guam and Puerto Rico used a simple random-sample design.

In the type of DSS design that states most commonly used in the BRFSS landline telephone sampling, BRFSS divides telephone numbers into two groups, or strata, which are sampled separately. The high-density and medium-density strata contain telephone numbers that are expected to belong mostly to households. Whether a telephone number goes into the high-density or medium-density stratum is determined by the number of listed household numbers in its hundred block, or set of 100 telephone numbers with the same area code, prefix, and first two digits of the suffix and all possible combinations of the last two digits. BRFSS puts numbers from hundred blocks with one or more listed household numbers ("1+ blocks," or "banks") in either the high-density stratum ("listed 1+ blocks") or medium-density stratum ("unlisted 1 + blocks"). BRFSS samples the two strata to obtain a probability sample of all households with telephones.

Cellular telephone sampling frames are commercially available and the system can call random samples of cellular telephone numbers, but doing so requires specific protocols. The basis of the 2013 BRFSS sampling frame is the Telecordia database of telephone exchanges (e.g., 617-492-0000 to 617-492-9999) and 1,000 banks (e.g., 617-492-0000 to 617-492-0999). The vendor uses dedicated cellular 1,000 banks, sorted on the basis of area code and exchange within a state. BRFSS forms an interval,  $K$ , by dividing the population count of telephone numbers in the frame,  $N$ , by the desired sample size,  $n$ . BRFSS divides the frame of telephone numbers into  $n$  intervals of size  $K$  telephone numbers. From each interval, BRFSS draws one 10-digit telephone number at random.

The target population for cellular telephone samples in 2013 consists of persons residing in a private residence or college housing, who have a working cellular telephone, are aged 18 and older, and received 90 percent or more of their calls on cellular telephones.

In the sample design, each state begins with a single stratum. To provide adequate sample sizes for smaller geographically defined populations of interest, however, many states sample disproportionately from strata that correspond to sub-state regions. In 2013, the 48 states or territories with disproportionately sampled geographic strata were Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts,

Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Pennsylvania, Puerto Rico, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Vermont, Virginia, the US Virgin Islands, Washington, and Wisconsin. As a precaution to protect the confidential responses provided by the respondent, specific variables (such as sub-state geographic identifiers, detailed race/ethnicity, and age greater than 80) in a given year are removed.

State health departments may directly collect data from their states or they may use a contractor. In 2013, 11 state health departments collected their data in-house; 42 contracted data collection to university survey research centers or commercial firms. In 2013, the CDC provided samples purchased from Marketing Systems Group, Inc. (MSG) to all 53 states and territories.

## **Data Collection**

### **Interviewing Procedures**

In 2013, 53 states or territories used Computer-Assisted Telephone Interview (CATI) systems. The CDC supports CATI programming using the Ci3 WinCATI software package. This support includes programming the core and module questions for data collectors, providing questionnaire scripting of state-added questions for states requiring such assistance, and contracting with a Ci3 consultant to assist states. Following guidelines provided by the BRFSS, state health personnel or contractors conduct interviews. The core portion of the questionnaire lasts an average of 18 minutes. Interview time for modules and state-added questions is dependent upon the number of questions used, but generally, they add 5 to 10 minutes to the interview.

Interviewer retention is very high among states that conduct the survey in-house. The state coordinator or interviewer supervisor conduct repeated training specific to the BRFSS. Contractors typically use interviewers who have experience conducting telephone surveys, but these interviewers are given additional training on the BRFSS questionnaire and procedures before they are approved to work on BRFSS.

BRFSS protocols require evaluation of interviewer performance. In 2013, all BRFSS surveillance sites had the capability to monitor their interviewers. Interviewer-monitoring systems vary from listening to the interviewer only at an on-site location to listening to both the interviewer and respondent at remote locations. Some states also use verification callbacks in addition to direct monitoring. Contractors typically conducted systematic monitoring of each interviewer a certain amount of time each month. All states had the capability to tabulate disposition code frequencies by interviewer. These data were the primary means for quantifying interviewer performance.

States conducted telephone interviews during each calendar month; they made calls seven days per week, during both daytime and evening hours. They followed standard BRFSS procedures for rotation of calls over days of the week and time of day. Detailed information on interview response rates is available in the BRFSS 2013 Summary Data Quality Report.

## Data Processing

### Preparing for Data Collection and Data Processing

Data processing is an integral part of any survey. Because states collect and submit data to the CDC each month of the year, BRFSS performs routine data processing tasks on an ongoing basis. And, once the final version of the new questionnaire becomes available each year, CDC staff takes steps to prepare for the next cycles of data collection. These steps include developing edit specifications, programming portions of the Ci3 WinCATI software, programming the editing software, and producing telephone sample estimates as requested by states and ordering the sample from the contract vendor. The CDC produces a Ci3 WinCATI data entry module for each state that requests it. CDC staff also must incorporate skip patterns, together with some consistency edits, and response-code range checks into the CATI system. These edits and skip patterns serve to reduce interviewer, data entry, and skip errors. Developers prepare data conversion tables that help processors read the survey data from the entry module, call information from the sample tracking module, and combine information into the final format for that data year. The CDC also creates and distributes a Windows-based editing program that can perform data validations on files with proper survey result formats. This program helps users with output lists of errors or warns users about conditions of concern that may exist in the data.

The CDC begins to process data for the survey year as soon as states or their contractors begin submitting data to the data management mailbox. Data processing continues throughout the survey year. The CDC receives and tracks monthly data submissions from the states. Once data are received from a state, CDC staff run editing programs and cumulative data quality checks and note any problems in the files. A CDC programmer works with each state until any problems are optimally resolved. CDC staff generate data quality reports and share them with state coordinators, who review the reports and discuss any potential problems. Once the CDC receives and validates the entire year of data for a state, processors run several year-end programs on the data. These programs perform some additional, limited data cleanup and fixes specific to each state and data year and produce reports that identify potential analytic problems with the data set. Once this step is completed, data are ready for assigning weights and adding calculated variables. Calculated variables are created for the benefit of users and can be noted in the data set by the leading underscore in the variable name. The following calculated variables are examples of results from this procedure:

- `_RFSMOK3`,



- `_MRACE1`,
- `_AGEG5YR`, and
- `_TOTINDA`.

For more information, see the [Calculated Variables and Risk Factors in Data Files](#) document. Several variables from the data file are used to create these variables in a process that varies in complexity; some are based only on combined codes, while others require sorting and combining of particular codes from multiple variables.

Almost every variable derived from the BRFSS interview has a code category labeled “refused”; which is assigned a value of “9,” “99,” or “999”. These values may also be used to represent missing responses. Missing responses may be due to non-interviews (a “non-interview” response results when an interview ends prior to this question and an interviewer codes the remaining responses as “refused”) and missing responses due to skip patterns in the questionnaire. However, this code may capture some questions that were supposed to have answers, but for some reason do not have them, and appeared as a blank or another symbol. Combining these types of responses into a single code requires vigilance on the part of data file users who wish to separate (1) results of respondents who did not receive a particular question and (2) results from respondents who, after receiving the question, gave an unclear answer or refused to answer it.

### **Weighting the Data**

When data are unweighted, each record counts the same as any other record. Unweighted data analyses make the assumption that each record has an equal probability of being selected and that noncoverage and nonresponse are equal among all segments of the population. When deviations from these assumptions are large enough to affect the results from a data set, weighting each record appropriately can help to adjust for assumption violations. In the BRFSS, such weighting serves as a blanket adjustment for noncoverage and nonresponse and forces the total number of cases to equal population estimates for each geographic region, which for the BRFSS sums to the state population. Regardless of state sample design, use of the final weight in analysis is necessary if users are to make generalizations from the sample to the population.

Following is a general description of the 2013 BRFSS weighting process. Where a factor does not apply, processors set its value to one for calculation. In order to reduce bias due to unequal probability of selection, design weighting is conducted. The BRFSS also uses iterative proportional fitting, or “raking” to adjust for demographic differences between those persons who are sampled and the population that they represent. Therefore the weighting methodology is comprised of two sections: design weight and raking.

Design weights are calculated using the weight of each geographic stratum (STRWT) the number of landline phones within a household (NUMPHON2) and the number of adults who use those phones (NUMADULT). For cellphone respondents, both NUMPHON2 and NUMADULT are set to 1. The formula for the design weight is:

$$\text{Design Weight} = \text{STRWT} * (1/\text{NUMPHON2}) * \text{NUMADULT}$$

In 2013, the inclusion of cellular telephone respondents who received between 90 and 99 percent of their telephone calls on their cellular telephone required an adjustment to the design weights to account for the overlapping sample frames. From each of the two sample frames, a compositing factor was calculated for the mostly cellular telephone dual sampling frame users. BRFSS multiplied the design weight by the compositing factor to generate a composite weight for the records in the overlapping sample frames. BRFSS then truncated the design weight based on quartiles within geographic region, which processors used as the raking input weight.

The stratum weight (STRWT) accounts for differences in the probability of selection among strata (subsets of area code/prefix combinations). It is the inverse of the sampling fraction of each stratum. There is rarely a complete correspondence between strata, defined by subsets of area code/prefix combinations, and regions, defined by the boundaries of government entities.

BRFSS calculates the stratum weight (STRWT) using the following items:

- *Number of available records* (NRECSTR) and the number of records users select (NRECSEL) within each geographic strata and density strata.
- *Geographic strata* (GEOSTR), which may be the entire state or a geographic subset such as counties, census tracts, etc.
- *Density strata* (\_DENSTR) indicating the density of the phone numbers for a given block of numbers as listed or not listed.

Within each \_GEOSTR\*\_DENSTR combination, BRFSS calculates the stratum weight (\_STRWT) from the average of the NRECSTR and the sum of all sample records used to produce the NRECSEL. The stratum weight is equal to NRECSTR / NRECSEL.

|             |   |
|-------------|---|
| 1/ NUMPHON2 | The inverse of the number of residential telephone numbers in the respondent's household. |
| NUMADULT    | The number of adults 18 years and older in the respondent's household.                    |

**FINAL WEIGHT** BRFSS rakes the design weight to 8 margins (age group by gender, race/ethnicity, education, marital status, tenure, gender by race/ethnicity, age group by race/ethnicity, phone ownership). If BRFSS includes geographic regions, it includes four additional margins (region, region by age group, region by gender, region by race/ethnicity). If at least one county has 500 or more respondents, BRFSS includes four additional margins (county, county by age group, county by gender, county by race/ethnicity).

**\_LLCPWT** The final weight assigned to each respondent.

BRFSS uses weight trimming to increase the value of extremely low weights and decrease the value of extremely high weights. The objective of weight trimming is to reduce errors in the outcome estimates caused by unusually high or low weights in some categories.

### *Calculation of a Child Weight*

BRFSS calculates the design weight for child weighting from the stratum weight times the inverse of the number of telephones in the household and then multiplies by the number of children:

$$\text{Child Design Weight} = \text{STRWT} * (1/\text{NUMPHON2}) * \text{CHILDREN}$$

**CHILDWT** = BRFSS rakes the child design weight to 5 margins including age by gender, race/ethnicity, gender by race/ethnicity, age by race/ethnicity, and phone ownership.

**\_CLLCPWT** is the weight assigned for each child interview.

## References

1. Remington PL, Smith MY, Williamson DF, Anda RF, et al. Design, characteristics, and usefulness of state-based behavioral risk factor surveillance: 1981–1987. *Public Health Reports* 1988;103(4):366–375.
2. US Census Bureau, 2008-2012 American Community Survey, Table B25043–Tenure by Telephone Service Available by Age of Householder-Universe: Occupied housing units, 2008–2012 American Community Survey 5-year estimates.  
<http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk>  
  
(Accessed June 10, 2014)
3. Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, January–June 2013. National Center for Health Statistics. December 2013. Available from <http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201312.pdf>
4. Battaglia MP, Frankel MR, Link MW. Improving standard poststratification techniques for random-digit–dialing telephone surveys. *Survey Research Methods* 2008;2:11–9. Available from <https://ojs.ub.uni-konstanz.de/srm/article/viewFile/597/1295>