



Attachment 7: BRFSS 2023 Weighting Documentation
and Comparability Technical Documentation

Contents

- BRFSS Overview and Weighting Documentation..... 3
 - Background..... 3
- The BRFSS Design..... 4
 - The BRFSS Questionnaire..... 4
 - Data Collection..... 7
 - Data Processing..... 8
- Weighting the Data..... 9
 - Calculation of a Child Weight..... 10
- References..... 10
- Comparability Documentation for 2023 BRFSS Data..... 12
 - 2023 Data Anomalies and Deviations from the Sampling Frame..... 12
 - Protocol Changes from 2023 Data Collection..... 12
 - Cellular Telephone Data..... 12
 - Weighting Methodologies..... 13
 - Other Issues..... 13
- Statistical and Analytic Issues..... 15
 - Analysis Procedures..... 15
 - Statistical Issues..... 15
 - Analytic Issues..... 16
- References..... 18

BRFSS Overview and Weighting Documentation

Background

The Behavioral Risk Factor Surveillance System (BRFSS) is a collaborative project between all the states in the United States 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 CDC's National Center for Chronic Disease Prevention and Health Promotion. The BRFSS is a system of ongoing health-related telephone surveys that collect data on health-related risk behaviors, chronic health conditions, health-care access, and use of preventive services from the noninstitutionalized adult population (≥ 18 years) residing in the United States and participating areas.

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; BRFSS now collects data in all 50 states as well as the District of Columbia and participating US territories. During 2023, 48 states, the District of Columbia, Guam, Puerto Rico, and the US Virgin Islands collected BRFSS data. Kentucky and Pennsylvania were unable to collect enough data to meet the minimum requirements to be included in the 2023 public data set.

In this document, the term "state" is used to refer to all areas participating in the BRFSS, including the District of Columbia, Guam, the Commonwealth of Puerto Rico, and the US Virgin Islands.

BRFSS's objective is to collect uniform state-specific data on health risk behaviors, chronic diseases and conditions, access to health care, and use of preventive health services related to the leading causes of death and disability in the United States. Factors assessed by the BRFSS in 2023 included health status and healthy days, exercise, hypertension, cholesterol, chronic health conditions, demographics, disability, falls, tobacco use, alcohol consumption, immunization, HIV/AIDS, seat belt use/drinking and driving, and long-term COVID effects (all core sections). Optional Module topics for 2023 included social determinants and health equity, reactions to race, prediabetes and diabetes, cognitive decline, caregiver, screenings of different types of cancer, cancer survivorship (type, treatment, pain management) and sexual orientation/gender identity (SOGI). Refer to the **2023 BRFSS Questionnaire** for the full list of topics and offerings.

Since 2011, the BRFSS has been conducting both landline telephone- and cellular telephone-based surveys. All the responses were self-reported; proxy interviews are not conducted by the BRFSS. In conducting the landline telephone survey, interviewers collected data from a randomly selected adult in a household. In conducting the cellular telephone survey, interviewers collected data from adults answering the cellular telephones residing in a private residence or college housing. Beginning in 2014, all adults contacted through their cellular telephone were eligible, regardless of their landline phone use (i.e., complete overlap).

The 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 use contractors. The data are

transmitted to CDC for editing, processing, weighting, and analysis. An edited and weighted data file is provided to each participating state health department for each year of data collection, and summary reports of state-specific data are prepared by CDC. State health departments use the BRFSS data for a variety of purposes, including identifying demographic variations in health-related behaviors; designing, implementing, and evaluating public health programs; addressing emergent and critical health issues; proposing legislation for health initiatives; and measuring progress toward state health objectives.¹

Health characteristics estimated from the BRFSS pertain to the noninstitutionalized adult population—aged 18 years or older—who reside in the United States. In 2023, an optional module was included to provide a measure for asthma prevalence for people aged 17 years or younger. BRFSS respondents are identified through telephone-based methods. According to the 2022 American Community Survey (ACS), 99.1% of all occupied housing units in the United States had telephone service available, and telephone non-coverage ranged from 1.0% or less in several states to 1.5% in Montana.² An estimated 2.5% of occupied households in Puerto Rico did not have telephone service.² 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. The preliminary results (July to December 2023) from the National Health Interview Survey (NHIS) indicate that 75.2% of adults were wireless-only.³ Using a dual-frame survey including landline and cellular telephones improved the validity, data quality, and representativeness of BRFSS data.

In 2011, a new weighting methodology called iterative proportional fitting (or “raking”)⁴ replaced the poststratification method to weight BRFSS data. Raking allows incorporation of cellular telephone survey data and permits the introduction of additional demographic characteristics (e.g., education level, marital status, home renter/owner) in addition to age, race/ethnicity and gender. These additional characteristics improve the degree and extent to which the BRFSS sample properly reflects the sociodemographic makeup of individual states. The 2023 BRFSS raking method includes categories of age by gender, detailed race and ethnicity groups, education levels, marital status, regions within states, gender by race or ethnicity, telephone source, renter or owner status, and age by race or ethnicity. In 2023, 48 states, the District of Columbia, Guam, Puerto Rico, and the US Virgin Islands collected samples of interviews conducted by landline and cellular telephone. (Kentucky and Pennsylvania did not collect enough data to be included in the 2023 public data set.)

The BRFSS Design

The BRFSS Questionnaire

Each year, the states—represented by their BRFSS coordinators and CDC—agree on the content of the questionnaire. The BRFSS questionnaire consists of a core component, optional modules, and state-added questions. 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 the BRFSS to take advantage of questions that have been tested and allows states to compare their data with those from other surveys. Any new core or module questions that states, federal agencies, or other entities propose as additions to the 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. The 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 as state-added questions.

The questionnaire has three parts:

1. 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, HIV/AIDS risks), as well as demographic questions. The core component includes the annual core-comprising questions asked each year and rotating core questions that are asked in even- and odd-numbered years.

2. Optional BRFSS modules: These are sets of questions on specific topics (e.g., social determinants of health and health equity, reactions to race, prediabetes and diabetes, cognitive decline, caregiver, screenings of different types of cancer, cancer survivorship and SOGI) that states elect to use on their questionnaires. 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 website.

3. State-added questions: Individual states develop or acquire these questions and add them to their BRFSS questionnaires. CDC does not edit, evaluate, track, or report responses from these questions.

The BRFSS supported 32 optional modules in 2023, but states limited their use of modules and state-added questions to only the most helpful to their state program purposes, in order to keep surveys at a reasonable length. Because different states have different needs, question totals vary widely between states. The 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 2023 list of optional modules used on both the landline telephone and cellular telephone surveys is available on the BRFSS website, published with the 2023 data materials. To allow for a wider range of questions in optional modules, combined landline telephone and cellular telephone data for 2023 include up to three split versions of the questionnaire. A split version is used 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 ***Modules Used by State and Modules Used by Category tables***, published with this yearly release.

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. Before the beginning of the calendar year, 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 then the state-added questions. This content order ensures comparability across states and follows the BRFSS guidelines. Generally, the only changes that the standard protocol allows are limited insertions of state-added questions on topics related to core questions. CDC and state partners must agree to these

exceptions. In some cases, however, states have not been able to follow all set guidelines. Users should refer to the yearly **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 CDC. States use the questionnaire without changes for one calendar year, and CDC archives a copy on the BRFSS website. 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, 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 2023. Forty-nine projects used a disproportionate stratified sample (DSS) design for their landline samples. Guam, Puerto Rico, and the US Virgin Islands used a simple random-sample design.

In the type of DSS design that states most commonly used in the BRFSS landline telephone sampling, the 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 2 digits of the suffix and all possible combinations of the last 2 digits. BRFSS puts numbers from hundred blocks with 1 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 2023 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. The BRFSS forms an interval— K —by dividing the population count of telephone numbers in the frame— N —by the desired sample size— n . The BRFSS divides the frame of telephone numbers into n intervals of size K telephone numbers. From each interval, the BRFSS draws one 10-digit telephone number at random.

The target population (aged 18 years and older) for cellular telephone samples in 2023 consists of people residing in a private residence or college housing who have a working cellular telephone.

In the sample design, states begin 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 2023, the 45 states with geographic stratification were:

Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Mexico, New York, North Carolina, Ohio, Oklahoma, Puerto Rico, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, US Virgin Islands, Utah, Vermont, Virginia, 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 or ethnicity, and older than 80 years) in a given year are removed.

State health departments may directly collect data from their state residents, or they may use a contractor. In 2023, 8 state health departments collected their data in-house and the remainder contracted with other data collectors. In 2023, the CDC provided samples purchased from Marketing Systems Group, Inc. (MSG) to all 52 states and territories.

Data Collection

Interviewing Procedures

In 2023, 54 states or territories used Computer-Assisted Telephone Interview (CATI) systems. (In this group, PA and KY were not able to collect enough data to be included in the 2023 data set.) 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 17 minutes. Interview time for modules and state-added questions depends on 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 conducts 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 the BRFSS.

The BRFSS protocols require evaluation of interviewer performance. During 2023, 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 7 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 2023 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 CDC each month, the BRFSS performs routine data processing tasks on an ongoing basis. Once the final version of the new questionnaire becomes available each year, CDC staff take 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, producing telephone sample estimates as requested by states and ordering the sample from the contract vendor. 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. 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.

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. 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 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
- _TOTINDA
- _HCVU651
- _AGE80
- _FLUSHOT7

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 and assigned values 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 then codes the remaining responses as refused.) and missing responses due to skip patterns in the questionnaire. This code, however, 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

The BRFSS is designed to obtain sample information on the population of interest i.e., the adult US population residing in different states. Data weighting helps make sample data more representative of the population from which the data were collected. BRFSS data weights incorporate the design of BRFSS survey and characteristics of the population. BRFSS weighting methodology comprises 1) design factors or design weight, and 2) some form of demographic adjustment of the population—by iterative proportional fitting or raking.

The design weight accounts for the probability of selection and adjusts for nonresponse bias and non-coverage errors. Design weights are calculated using the weight of each geographic stratum (`_STRWT`), the number of phones within a household (`NUMPHON3`), and the number of adults aged 18 years and older in the respondent's household (`NUMADULT`). For cellphone respondents, both `NUMPHON3` and `NUMADULT` are set to 1. The formula for the design weight is

$$\text{Design Weight} = _STRWT * (1/NUMPHON3) * NUMADULT$$

The stratum weight (`_STRWT`) accounts for differences in the probability of selection among strata (subsets of area code or prefix combinations). It is the inverse of the sampling fraction of each stratum. There is rarely a complete correspondence between strata (which are defined by subsets of area code or prefix combinations) and regions—which are 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 (e.g., counties, census tracts).
- 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`.

The complete overlapping sample frames required an adjustment to address the respondent's probability of selection in both the landline and cell phone sample frame. A compositing factor was calculated for dual users in landline and cell phone sample frames. The design weight is adjusted by the compositing factor for the records in the overlapping sample frames and later truncated within

geographic region using the mean ± 1.96 times the standard deviation to calculate the truncation limits. The adjusted and truncated design weight was used as the raking input weight.

BRFSS uses iterative proportional fitting, or raking, to adjust for demographic differences between those persons who are sampled and the population that they represent. After combining landline and cellular telephone data, BRFSS performs raking by adjusting one or a combination of demographic categories at a time in an iterative process until a convergence of a set value is reached. The BRFSS rakes the design weight to 8 margins (gender by age group, race or ethnicity, education, marital status, tenure, gender by race or ethnicity, age group by race or ethnicity, and phone ownership). If the state had geographic regions, it includes 4 additional margins (region, region by age group, region by gender, and region by race or ethnicity). If the state had at least 1 county with 500 or more respondents, the BRFSS includes 4 additional margins (county, county by age group, county by gender, and county by race or ethnicity). BRFSS, therefore, uses the adjusted and truncated design weight for raking and produces `_LLCPWT`—the final weight assigned to each respondent.

The population estimates obtained for building the target totals for raking are from similar sources used in previous years. Postcensal population estimates were purchased from Claritas, LLC at the county-level for age, race or ethnicity, and gender. These population estimates are used as the population totals for a state across all margins. The 5-year American Community Survey PUMS data set (2018–2022) was used to obtain estimates for margins 3, 4, and 5 (education, marital status, tenure). The noninstitutionalized adults were weighted by the person-level weights to generate the population estimates. The percentages were then used in the raking margins. The telephone ownership estimates for margin 8 were taken from the state wireless estimate percentages produced by the National Center for Health Statistics (NCHS) and released in December of 2022.

Calculation of a Child Weight

The 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:

$$\text{Child Design Weight} = \text{_STRWT} * (1/\text{NUMPHON3}) * \text{CHILDREN}$$

`CHILDDWT` = BRFSS rakes the child design weight to 5 margins including age by gender, race or ethnicity, gender by race or ethnicity, age by race or ethnicity, and phone ownership.

`_CLLCPWT` is the weight assigned for each child interview.

References

1. Remington PL, Smith MY, Williamson DF, Anda RF, Gentry EM, Hogelin GC. Design, characteristics, and usefulness of state-based behavioral risk factor surveillance: 1981-87. *Public Health Rep.* 1988;103(4):366-375.
2. Federal Communications Commission USA. Universal Service Monitoring Report. 2023; <https://www.fcc.gov/document/2023-universal-service-monitoring-report> . Table 6.6 (p63) Accessed August 21st, 2024.
3. Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, July–December 2023. National Center for Health Statistics. June 2024. Available from:

<https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless2ken02406.pdf> . Accessed August 27th, 2024.

4. MP B. Improving standard poststratification techniques for random-digit-dialing telephone surveys. *Surv Res Methods*. 2008;2(1):9.

Comparability Documentation for 2023 BRFSS Data

2023 Data Anomalies and Deviations from the Sampling Frame

The BRFSS state-based annual sample designs are fixed for the data collection year beginning in January in all the states and participating US territories. The samples are drawn quarterly and screened monthly to provide a representative sample for monthly data collection. The intent of the monthly sample is to use it for 1 month, but in most states, it took more than 1 month to complete data collection using the monthly sample. In several instances, states used their monthly sample during a period of several months. This deviation will disproportionately affect analyses based on monthly (rather than annual) data.

Several states conducted fewer than 12 monthly telephone samples for data collection during the year. The following states did not collect 12 monthly **landline** samples: Alabama, Arizona, Connecticut, Idaho, Indiana, Kansas, Louisiana, Minnesota, Mississippi, Nevada, New Hampshire, New Mexico, North Carolina, North Dakota, South Carolina, South Dakota, Tennessee, Virginia, District of Columbia, Guam, Puerto Rico, and the US Virgin Islands.

The following states did not collect 12 monthly **cellphone** samples: Alabama, Arizona, Connecticut, Idaho, Indiana, Kansas, Louisiana, Maine, Minnesota, Mississippi, Nevada, New Hampshire, New Mexico, North Carolina, North Dakota, Oregon, South Carolina, South Dakota, Tennessee, West Virginia, the District of Columbia, Guam, Puerto Rico, and the US Virgin Islands.

Thirty-four states, Guam, Puerto Rico, and the US Virgin Islands were unable to close-out their 2023 sample by December 31, 2023, and continued data collection into early 2024.

Arizona, Idaho, Indiana, and the District of Columbia began data collection in February. Minnesota began data collection in March. Idaho, Indiana, and South Carolina began data collection in May. North Dakota began data collection in June.

The months of data collection missed in each situation will likely affect seasonal estimates, i.e. influenza vaccination. Forty-eight states, the District of Columbia, Guam, Puerto Rico, and the US Virgin Islands met the minimum requirements to be included in the public-use data set for 2023; please consider the differences in collection when comparing estimates across years.

Protocol Changes from 2023 Data Collection Cellular Telephone Data

Telephone coverage varies by state and also by subpopulation. According to the 2022 American Community Survey (ACS), 99.1% of all occupied housing units in the United States had telephone service available; telephone non-coverage ranged from 1.0% or less in several states to 1.5% in Montana.³ It is estimated that 2.5% of occupied households in Puerto Rico did not have telephone service.³ The increasing percentage of households abandoning their landline telephones for cellular telephones has significantly eroded the population coverage provided by landline telephone-based surveys to pre-1970s

levels. The preliminary results (July to December 2023) from the National Health Interview Survey (NHIS) indicate that 75.2% of adults were wireless-only.⁴ The increased use of cellular telephones required the BRFSS to begin to include the population of cellular telephone users in 2011. At that time, all adult cellular telephone respondents who had a landline telephone were not eligible for the survey. In 2012, the BRFSS changed the screening process. Cellular telephone respondents were eligible—even if they had landline phones—as long as they received at least 90% of all calls on their cell phones. Beginning in 2014, all adults contacted through their personal (nonbusiness) phone numbers were eligible regardless of their landline phone use (i.e., complete overlap).

Weighting Methodologies

Since 2011, BRFSS has used the weighting methodology called iterative proportional fitting (IPF)—or raking—to weight data. Raking allows incorporation of cellular telephone survey data, and it permits the introduction of additional demographic characteristics that more-accurately match sample distributions to known demographic characteristics of populations at the state level. (Refer to the CDC website for more information on [methodologic changes](#)). Raking adjusts the estimates within each state using the margins (raking control variables). The raking method applies a proportional adjustment to the weights of the cases that belong to the same category of the margin. The iteration (up to 100 times) continues until a convergence to within a target percentage difference is achieved. Since 2013, up to 16 raking margins have been used in the following order—county by gender, county by age, county by race or ethnicity, county, region by race or ethnicity, region by gender, region by age, region, telephone service (landline, cellular telephone or dual user), age by race or ethnicity, gender by race or ethnicity, tenure (rent or own), marital status, education, race or ethnicity, and gender by age.

Since 2014, the inclusion of all adult cellular telephone respondents in the survey required an adjustment to the design weights to account for the overlapping sample frames. A compositing factor was calculated from each of the two samples (landline and cellular sample) for dual users—individuals who had both cellular telephone and landline phone. The BRFSS multiplied the design weight by the compositing factor to generate a composite weight for the records in the overlapping sample frame. Later the design weight was truncated based on quartiles within geographic region (or state). In 2023, the truncated weight was adjusted to regional (or state) population and the state phone source proportions prior to raking. This adjusted weight was used as the input weight for the first raking margin. At the last step of the raking process, weight trimming was used to increase the value of extremely low weights and decrease the value of extremely high weights. Weight trimming is based on two alternative methods, IGCV (Individual and Global Cap Value) and MCV (Margin Cap Value).

Other Issues

As in previous years, the data from an optional module were included if interviewers asked module questions to all eligible respondents within a state for the entire data collection year. A state may have indicated the use of an optional module. If the module was not used for the entire data collection year, the data were moved into the state-added questions section. Several states collected data with optional modules by landline telephone and cellular telephone surveys.

CDC has also provided limited technical support for the survey data collection of multiple (up to three in 2023) questionnaire versions. A state may ask a subset of its survey sample a different set of questions following the core, as long as the survey meets the minimum effective sample size (2,500 participants) for a given questionnaire version. States must use the core instrument without making any changes to it in any of their versions of the overall questionnaire. States can include an optional module on all versions or exclusively on a single version but, once a state chooses to use an optional module, the state must ask the module questions throughout the data collection year. The objective of the multiple-version questionnaire is to ask more questions, on additional topics, within a statewide sample. In 2023, 11 states conducted multiple-questionnaire-version surveys on both their landline telephone and cellular telephone surveys. Data users can find version-specific data sets and additional documentation regarding module data analysis in the 2023 BRFSS Survey Data and Documentation.

A 2012 change to the final disposition code assignment rules modified the requirements for a partially complete interview. If a participant terminated an interview during or after the demographics section, the BRFSS coded it as a partial-complete. The coding of questions was discontinued at the point of interview termination. When determining which records to include in any analysis, data users should account for participants' missing and refused values. Beginning in 2015, questions in the demographic section were reordered and the definition of a partial-complete changed. A partially complete disposition code in 2023 was assigned if the interview terminated before completion of the survey and the selected respondent completed the demographics section through question 12 for a cell phone interview and for a landline interview.

More information about survey item nonresponse can be found in the 2023 BRFSS Summary Data Quality Report and in the respective states' Data Quality Reports.

Alabama continued 2023 data collection into 2024 using the state's January 2024 sample to collect an additional month of data with its 2023 survey. The sample design did not change across years. The effort to collect additional interviews was to make up for interviews missed during the first two months of 2023 data collection.

Maryland's 2023 data collector did not collect the SHINGLE2 question from cell phone respondents, which identified as out-of-state residents. The error in the CATI program skips logic for this question during the month of January. This issue was corrected while February and March data collection were in

the field. There are missing responses from the January, February and March data in some cell phone records transferred to another state.

Mississippi's 2023 data collector incorrectly coded a skip instruction for two optional modules (Heart Attack and Stroke, Aspirin for CVD Prevention) during the first quarter of the 2023 data collection. The skip excluded respondents who were not employed full or part time. The responses remain blank for the two modules where the data collector was not able to obtain answers from the respondent.

Statistical and Analytic Issues

Analysis Procedures

To use the BRFSS data, the researcher needs to formulate a research question, review the existing data tabulations, develop an analytic plan, conduct the analyses, and use data for decision making.⁵ Unweighted BRFSS data represent the actual responses of each respondent before any adjustment is made for variation in the respondents' probability of selection, disproportionate selection of population subgroups relative to the state's population distribution, or nonresponse. Weighted BRFSS data represent results that have been adjusted to compensate for these issues. Regardless of state sample design, use of the weight in analysis is necessary if generalizations are to be made from the sample to the population. Please note the statistical and analytic issues described in this section are the same as those of previous years.

Statistical Issues

The procedures for estimating variances described in most statistical texts and used in most statistical software packages are based on the assumption of simple random sampling (SRS). The data collected in the BRFSS, however, are obtained through a complex sample design; therefore, the direct application of standard statistical analysis methods for variance estimation and hypothesis testing may yield misleading results. There are computer programs available that take such complex sample designs into account: SAS Version 9.4 SURVEYMEANS and SURVEYREG procedures, SUDAAN, and Epi Info's C-Sample are among those suitable for analyzing BRFSS data.^{6,7,8} SAS and SUDAAN can be used for tabular and regression analyses.^{6,7} Epi Info's C-sample can be used to calculate simple frequencies and two-way cross-tabulations.⁸ When using these software products, users must know the stratum, the primary sampling units, and the record weight—all of which are on the public use data file. For more information on calculating variance estimations using SAS, see the SAS/STAT® 13.1 User's Guide.⁶ For information about SUDAAN, see the SUDAAN Language Manual, Release 117, and to find more about Epi Info, see *Epi Info, Version 7.0.8*

Although the overall number of respondents in the BRFSS is more than sufficiently large for statistical inference purposes, subgroup analyses can lead to estimates that are unreliable. Consequently, users

need to pay particular attention to the subgroup sample when analyzing subgroup data, especially within a single data year or geographic area.

Small sample sizes may produce unstable estimates. Reliability of an estimate depends on the actual unweighted number of respondents in a category, not on the weighted number. Interpreting and reporting weighted numbers based on a small, unweighted number of respondents can mislead the reader into believing that a given finding is much more precise than it actually is. The BRFSS previously followed a rule of not reporting or interpreting percentages based upon a denominator of fewer than 50 respondents (unweighted sample) or the half-width of a 95% confidence interval greater than 10.

From 2011, the BRFSS replaced the confidence interval limitation with the relative standard error (RSE) —the standard error divided by the mean. The survey with the lower RSE has a more-precise measurement. Because there is less variance around the mean, BRFSS did not report percentage estimates where the RSE was greater than 30% or the denominator represented fewer than 50 respondents from an unweighted sample. Details of changes beginning with the 2011 BRFSS are available in the *Morbidity and Mortality Weekly Report (MMWR)*, which highlights weighting and coverage effects on trend lines.⁹ Because of the changes in the methodology, researchers are advised to avoid comparing data collected before the changes (up to 2010) with data collected from 2011 and onward.

Analytic Issues

a. Advantages and Disadvantages of Telephone Surveys

Compared with face-to-face interviewing techniques, telephone interviews are easy to conduct and monitor and are cost-efficient; however, telephone interviews have limitations. Telephone surveys may have higher levels of no coverage than face-to-face interviews because interviewers may not be able to reach some US households by telephone. As mentioned earlier, approximately 99% of households in the United States have telephones.³ A number of studies have shown that the telephone and non-telephone populations are different with respect to demographic, economic, and health characteristics.^{10,11,12} Although the estimates of characteristics for the total population are unlikely to be substantially affected by the omission of the households without telephones, some of the subpopulation estimates could be biased. Telephone coverage is lower for population subgroups such as people with low incomes, people in rural areas, people with less than 12 years of education, people in poor health, and heads of households younger than 25 years of age.¹³ Raking adjustments for age, race, and sex, and more demographic variables, however, minimize the impact of differences to a greater extent in no coverage, under-coverage, and nonresponse at the state level.

Surveys based on self-reported information may be less-accurate than those based on physical measurements. For example, respondents are known to underreport body weight and risky health behaviors, such as alcohol intake and smoking. This type of potential bias arises when conducting both

telephone and face-to-face interviews and when interpreting self-reported data, data users should take into consideration the potential for underreporting.

Despite the above limitations, the BRFSS data are reliable and valid.¹⁴ The prevalence estimates from the BRFSS correspond well with findings from surveys based on face-to-face interviews, including the National Health Interview Survey (NHIS), and the National Health and Nutrition Examination Survey (NHANES).¹⁵ Please visit the BRFSS website for more information about [methodological studies](#).

b. New Calculated Variables and Risk Factors

Not all of the variables that appear on the public use data set are taken directly from the state files. CDC prepares a set of SAS programs that are used for end-of-year processing. These programs prepare the data for analysis and add weighting, sample design, calculated variables, and risk factors to the data set. The following calculated variables and risk factors, which the BRFSS has created for the user's convenience, are examples of results from this procedure for 2023 data:

`_TOTINDA, _PNEUMO3, _RFBING6, _RFSMOK3, _RFHLTH, _CASTHM1, _RFHYPE6`

The procedures for calculating the variables vary in complexity. Some only combine codes, while others require sorting and combining selected codes from multiple variables. This may result in the calculation of an intermediate variable. For more information regarding the calculated variables and risk factors, refer to the document entitled *Calculated Variables in the 2023 Data File of the Behavioral Risk Factor Surveillance System*, found in the 2023 BRFSS Survey Data and Documentation section of the BRFSS website.

Two calculated variables (`_METSTAT`, `_URBSTAT`) have been included based on the 2013 NCHS urban-rural classification scheme for counties.¹⁶ The two variables identify metropolitan status versus nonmetropolitan or urban versus rural within a given state. Three states had a single county in a nonmetropolitan or rural category, thus requiring a recode of the value to an adjacent category as a disclosure-avoidance measure. The definitions below show the categorization of the two variables based on the sub-setting of the original six categories.

`_METSTAT` :

1 = `_URBNRRL` IN (1,2,3,4) = Metropolitan counties

2 = `_URBNRRL` IN (5,6) = Nonmetropolitan counties

`_URBSTAT` :

1 = `_URBNRRL` IN (1,2,3,4,5) = Urban counties

2 = `_URBNRRL` IN (6) = Rural counties

References

1. Mokdad AH, Stroup DF, Giles WH. Public health surveillance for behavioral risk factors in a changing environment: recommendations from the Behavioral Risk Factor Surveillance team. *MMWR Recomm Rep*. 2003;52(RR-9):1-12.
2. Holtzman D. The Behavioral Risk Factor Surveillance System. In: Blumenthal DS, DiClemente RJ, eds. *Community-Based Health Research: Issues and Methods*. New York, NY: Springer Publishing Company Inc; 2004:115-131.
3. Federal Communications Commission USA. Universal Service Monitoring Report. 2023; [DOC-401168A1.pdf \(fcc.gov\)](#) page 63. Accessed August 2024.
4. Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, July–December 2023. National Center for Health Statistics. June 2024. Available from: [Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July-December 2023 \(cdc.gov\)](#) Accessed August 2024.
5. Frazier EL, Franks AL, Sanderson LM, Centers for Disease Control and Prevention. Behavioral risk factor data. In: *Using Chronic Disease Data: A Handbook for Public Health Practitioners*. Atlanta, GA: Centers for Disease Control and Prevention, US Dept. of Health and Human Resources; 1992.
6. SAS Institute Inc. 2013, SAS/STAT® 13.1 User's Guide. Cary, NC: SAS Institute, Inc.
7. Research Triangle Institute (2012). SUDAAN Language Manual, Vols 1 and 2, Release 11.
8. Dean AG, Arner TG, Sunki GG, et al. Epi Info™, a database and statistics program for public health professionals. Atlanta, GA: Centers for Disease Control, US Dept of Health and Human Resources; 2011.
9. Pierannunzi C, Town M, Garvin W, et al. Methodologic changes in the Behavioral Risk Factor Surveillance System in 2011 and potential effects on prevalence estimates. *MMWR Morb Mortal Wkly Rep*. 2012;61(22):410-413. www.cdc.gov/mmwr/preview/mmwrhtml/mm6122a3.htm Accessed 29 August 2024.

10. Groves RM, Kahn RL. *Surveys by Telephone: A National Comparison with Personal Interviews*, New York, NY: Academic Press; 1979.
11. Banks MJ. Comparing health and medical care estimates of the phone and nonphone populations. In: *Proceedings of the Section on Survey Research Methods*. American Statistical Association. 1983:569-574.
12. Thornberry OT, Massey JT. Trends in United States telephone coverage across time and subgroups. In: Groves RM, et al, eds. *Telephone Survey Methodology*. New York, NY: John Wiley & Sons; 1988:25-49.
13. Massey JT, Botman SL. Weighting adjustments for random digit dialed surveys. In: Groves RM, et al, eds. *Telephone Survey Methodology*. New York, NY: John Wiley & Sons; 1988:143-160.
14. Pierannunzi C, Hu SS, Balluz L. A systematic review of publications assessing reliability and validity of the Behavioral Risk Factor Surveillance System (BRFSS), 2004–2011. *BMC Med Res Methodol*. 2013;13:49.
15. Li C, Balluz L, Ford ES, et al. A comparison of prevalence estimates for selected health indicators and chronic diseases or conditions from the Behavioral Risk Factor Surveillance System, the National Health Interview Survey, and the National Health and Nutrition Examination Survey, 2007-2008. *Prev Med*. 2012;54(6):381-387.
16. Ingram DD, Franco SJ. 2013 NCHS Urban-Rural Classification Scheme for Counties. National Center for Health Statistics. *Vital Health Stat*. 2014;2(166):1-73.