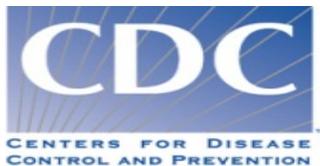


# Behavioral Risk Factor Surveillance System

Module Data for Analysis for 2013 BRFSS

August 2014





## Overview

The Behavioral Risk Factor Surveillance System (BRFSS) is a state-based telephone survey that includes information on a number of health outcomes, risk behaviors, use of preventive services, and chronic conditions for non-institutionalized adults who reside in each of the states and selected US territories. In 2011, BRFSS changed its data collection procedures, structure, and weighting methodology so the traditional landline telephone-based data set could include data from participants using cellular telephones as well. The BRFSS includes a core set of questions and multiple optional modules that focus on specific health issues. Not all modules are collected by all of the states, and states may opt to include module data on portions of their surveys. Core questions are included on all interviews. Researchers using BRFSS data should conduct analyses with complex sampling procedures; they also should appropriately stratify weighted the data in their work. . Weighting can improve selection probabilities and noncoverage among segments of the population.

Researchers conducting analysis of variables from the core-only section should use the variable `_LLCPWT` for weighting. BRFSS has created this document to guide users analyzing variables from 2013 BRFSS modules and/or combinations of module and core variables. As was true in 2012, the 2013 BRFSS data set includes split versions of module data that respondents provided by cellular telephone. Data users should note that new weighting procedures are likely to affect trend lines when comparing BRFSS data collected before and after 2011; because of these changes, users are advised NOT to make direct comparisons with pre-2011 data, and instead, begin new trend lines with that year.

**Data users should become familiar with the information presented in this document prior to performing analyses.** More information about the changes to the 2013 BRFSS is available at [http://wwwdev.cdc.gov/brfss/annual\\_data/2013/pdf/Compare\\_2013.pdf](http://wwwdev.cdc.gov/brfss/annual_data/2013/pdf/Compare_2013.pdf)

## Using BRFSS Data from Multiple Datasets

BRFSS has grouped the 2013 data into eight datasets:

- [2013 BRFSS questionnaire data \(combined landline telephone and cellular telephone\).](#)
- [2013 BRFSS combined landline telephone and cellular telephone multiple version questionnaire version 1 data \(landline telephone and cellular telephone version 1\).](#)
- [2013 BRFSS combined landline telephone and cellular telephone multiple version questionnaire version 2 data \(landline telephone and cellular telephone version 2\).](#)
- [2013 BRFSS combined landline telephone and cellular telephone multiple version questionnaire version 3 data \(landline telephone and cellular telephone version 3\).](#)
- [2013 BRFSS landline telephone data \(landline telephone only\)](#)
- [2013 BRFSS landline telephone multiple version questionnaire version 1 data \(landline telephone version 1\).](#)
- [2013 BRFSS landline telephone multiple version questionnaire version 2 data \(landline telephone version 2\).](#) and
- [2013 BRFSS landline telephone multiple version questionnaire version 3 data \(landline telephone version 3\).](#)

Individual states may have chosen to use a number of optional modules, depending on each state's needs. Although core questions are always collected on both landline telephone and cellular telephones, states may have collected module data by landline telephone and/or cellular telephone. As was true in 2012, some states also collected module data by combined landline telephone and cellular telephone in 2013. By splitting the modules, the states divided their samples and used different modules in the subsamples that were distinguished by the version of the surveys. Some modules may appear only on versions of questionnaires given to landline telephone respondents. Modules that appear in each version of a state's questionnaire are called common modules. Each state that collected common modules did so on all its versions of combined landline telephone and cellular telephone or landline telephone-only data.

Prior to using optional module data, users must identify the states that collected the data of interest and determine which questionnaire version the state used, if any. A list of modules, by state, is available here: <http://apps.nccd.cdc.gov/BRFSSModules/ModByState.asp?Yr=2013>.

As a first step to conducting research using the BRFSS module data, users should understand that the data set they need is based on the location of the questions in the core, in optional modules collected by landline telephone, or in optional modules collected by landline telephone and cellular telephone. Users should keep in mind that there are eight possible data sets from which they will need to include data:

- 1) Researchers use the combined landline telephone and cellular telephone data if the questions are exclusively from the core section or common modules asked on both the landline survey and cellular telephone survey.
- 2) Researchers use the landline telephone survey data if states had asked the questions only in the common version of the landline telephone survey.
- 3) Researchers use the combined landline telephone and cellular telephone data sets from states collecting multiple version questionnaires (up to 3) and landline telephone data sets from states collecting multiple version questionnaires (up to 3).

In all cases, researchers should use the variable `_STSTR` for stratification and `_PSU` should be used for cluster in complex sampling analyses. The following table displays the description of the data, name of the data sets, and the variable names of the final weight.

Data description	Data set name	Final weight variable name
Combined landline telephone and cellular telephone	LLCP2013	<code>_LLCPWT</code>
Combined landline telephone and cellular telephone Version 1	LLCP13V1	<code>_LCPWTV1</code>
Combined landline telephone and cellular telephone Version 2	LLCP13V2	<code>_LCPWTV2</code>
Combined landline telephone and cellular telephone Version 3	LLCP13V3	<code>_LCPWTV3</code>
Landline telephone only	LAND2013	<code>_LANDWT</code>
Landline telephone only Version 1	LAND13V1	<code>_LNDWTV1</code>
Landline telephone only Version 2	LAND13V2	<code>_LNDWTV2</code>
Landline telephone only Version 3	<code>_LNDWTV1</code>	<code>_LNDWTV3</code>

## Examples of Preparing 2013 BRFSS Module Data for Analysis

Due to the complex nature of state-based data collection processes, users may have to create a data set that fits their research needs. The following examples illustrate how to prepare module data for analysis when states have collected module data in a variety of ways.

Prior to the analysis, users should always go through all documents relative to the given data year before conducting analysis and double-check state FIPS codes in each data set to avoid duplication.

### Example 1: Colorectal Cancer Screening Module

The example below uses the Colorectal Cancer Screening Module to demonstrate how to combine and reweight data from multiple data sets for analysis. BRFSS selected this module because states collected the data either as a common module in the combined landline telephone and cellular telephone or

landline telephone-only data, or by splitting the sample and offering the module on one of the versions of the survey in the combined landline telephone and cellular telephone or landline telephone-only data.

**1. In the 2013 landline telephone and cellular telephone data, search for states that have chosen to use this module that is common:**

Go to the modules [Web page](#) to find states that have data for the Colorectal Cancer Screening module recorded in the combined landline telephone-cellular telephone data set. These states are Alabama (1\*), Arizona (4), Delaware (10), Florida (12), Georgia (13), Hawaii (15), Michigan (26), Nebraska (31), Nevada (32), New Hampshire (33), and Washington (53). (\*Note: Numbers in parentheses are the State FIPS codes, which can be found at <http://www.bls.gov/lau/lausfips.htm>.)

**2. Search for states that collected data for this module with questionnaire version(s) in the 2013 landline telephone and cellular telephone data:**

Go to the Web page for the [combined landline telephone and cellular telephone multiple version questionnaire data sets](#) and click on the document named “[2013 Combined Landline Telephone and Cellular Telephone Survey Multiple-Version Questionnaire Use of Data](#).” This document lists modules by state and states by module. Go to page 5 where the states are listed under the module of Colorectal Cancer Screening. California (6) and Oklahoma (40) are the two states that collected data on Version 1; New York (36) collected data on Version 2; New Jersey (34) collected data on Version 3 in the combined landline telephone and cellular telephone set.

**3. Search for states that collected data for this module among landline telephone-only users, common module:**

Go to the Web page for the [landline questionnaire data sets](#) and click on the document named “[Use of 2013 Land Line Questionnaire Data](#)”; Go to the table listed by modules, which begins on page 6. Search for “Colorectal Cancer Screening” in this table to find the states that used landline telephone data. These states listed include not only those in the combined landline telephone and cellular telephone data that are listed in step 1 above, but also common modules of landline telephone only. Compare this list with the list of states in step 1; you will find Massachusetts (25) is the state that used this common module in the landline telephone only data.

**4. Obtain data sets:**

- To download the landline telephone and cellular telephone data set (LLCP2013):  
Go to the [data Web page](#) and choose the 2013 data set. From there, please read all the 2013 documents including the [overview](#), [codebook](#), and the [module list](#). The data set is available in ASCII and SAS transport format files. A record layout, format syntax, and SAS code are included to read the raw data. The Web page also provides the link to access 2013 BRFSS landline telephone data or click on the Web links below to find the needed landline files.
- To download the landline telephone and cellular telephone dataset with questionnaire versions 1, 2, and 3:
  - Go the combined landline and cellular telephone multiple version questionnaire module data Web page to download multiple-version landline telephone data. There are 3

separate SAS datasets corresponding to version 1, version 2, and version 3. Choose Version 1: LLCP13V1, Version 2: LLCP13V2, and Version 3: LLCP13V3.

- To download the landline telephone-only dataset (LAND2013):
  - Go to the landline questionnaire data set Web page. The Web site also provides access to the 2013 landline telephone multiple version questionnaires.

Generate a new and uniform final weight variable from each of the data sets and combine data into a working data set. This step will require users to:

- a. keep all the states that collect the module in each of the data sets,
- b. rename the corresponding weight variable to a consistent weight variable,
- c. combine all the data sets into one data set that contains the consistent weight variable.

Details are shown in the SAS syntax on the next page.

## Example SAS code:

```
/*Example SAS code*/
*extract states from landline and cell phone combined data;
data llcp; *observation number=118,775;
  *originally downloaded data selecting states that collected data using
  Colorectal Cancer
  Screening Module from combined landline and cell phone data for common module;
  set in.llcp2013 (where=(state in (1,4,10,12,13,15,26,31,32,33,53)));
  *state FIPS code:
1(Alabama),4(Arizona),10(Delaware),12(Florida),13(Georgia),15(Hawaii),
  26(Michigan),31(Nebraska),32(Nevada),33(New Hampshire),53(Washington);
  _finalwt = _llcpwt;
  *rename final weight variable to be consistent across new data sets;
  drop _llcpwt;
run;

*extract states from combined landline and cell phone data, multiple versions;
*multiple version 1;
data llcpv1; *observation number=7,990;
  *originally downloaded data selecting states that collected data for the
  Colorectal Cancer
  Screening Module from combined landline and cell phone data, multiple versions;
  set in.llcp13v1 (where=(state in (6,40))); *state FIPS code: 6(California),
40(Oklahoma);
  *rename final weight variable to be consistent across new data sets;
  _finalwt=_lcpwtv1;
  drop _lcpwtv1;
run;

data llcpv2; *observation number=3,901;
  *originally downloaded data selecting states that collected data for the
  Colorectal Cancer
  Screening Module from combined landline and cell phone data, multiple versions;
  set in.llcp13v2 (where=(state=36)); *State FIPS code: 36(New York);
  *rename final weight variable to be consistent across new data sets;
  _finalwt=_lcpwtv2;
  drop _lcpwtv2;
run;

data llcpv3; *observation number=4,408;
  *originally downloaded data selecting states that collected data for the
  Colorectal Cancer
  Screening Module from combined landline and cell phone data, multiple versions;
  set in.llcp13v3 (where=(state=34)); *State FIPS code: 36(New Jersey);
  *rename final weight variable to be consistent across new data sets;
  _finalwt=_lcpwtv3;
  drop _lcpwtv3;
run;

*extract states from landline only data, common module;
data land; *observation number=12,160;
  *originally downloaded data, selecting state that collected data for the
  Colorectal Cancer
  Screening Module from landline only data, common module;
  set in.land2013 (where=(state=25)); *State FIPS code: 25(Massachusetts);
  *rename final weight variable to be consistent across new data sets;
  _finalwt = _landwt;
  drop _landwt;
```

```
run;

*combine all data sets together;
data colscreen; *observation number=147,234;
  set llcp llcpv1 llcpv2 llcpv3 land;
run;
```

## Example 2: Arthritis Management Module

The example below uses the Arthritis Management Module to demonstrate how to combine and reweight data from multiple data sets. This module was selected because states collected the data by splitting the sample and offering the module on one or more versions of the surveys in combined landline and cellular telephone or landline telephone-only data.

**1. Search for states that used this common module to collect data by landline telephone and cellular telephone:**

Go to the module [Web page](#) to search for this module. Seven states collected data as a common module by combined landline and cellular telephone: Kentucky (21), Minnesota (27), Mississippi (28), Missouri (29), Pennsylvania (42), Tennessee (47), and Utah (49).

**2. Search for states that used this module to collect data by landline telephone and cellular telephone in multiple versions:**

Go to the [combined landline telephone and cellular telephone multiple version questionnaire module data](#) Web page and click on the document titled, “[2013 Combined Landline Telephone and Cellular Telephone Survey Multiple-Version Questionnaire Use of Data.](#)” and find table listed by module. Under “Arthritis Management”, note that Michigan (26) collected data for this module in both version 2 and version 3; California (6), Kansas (20), Montana (30), New York (36), and Oregon (41) collected data for this module in version 2.

**3. Obtain datasets:**

To download the landline telephone and cellular telephone data set with multiple versions: Go to the [Combined Landline Telephone and Cellular Telephone Multiple Version Questionnaire Module Data](#) Web page. From there, please read all the support documents including the overview, codebooks, and the module lists available for each version. The data sets come in ASCII, SAS transport formats, and SAS files with record layout, format syntax to read the raw data. Download combined landline telephone and cellular telephone data sets of version 2 and version 3.

4. To generate a new and uniform final weight variable from each of the data sets and combine these data sets into one new data set for analysis, follow the SAS syntax, below. In some of the modules (e.g., Arthritis Management) where the states have multiple versions of combined landline telephone and cellular telephone data, run a frequency table to check the sample size in each of the multiple versions of the data. Compare the sample size in these versions and decide how to calculate the final weight by multiplying by a proportion of the whole. Here is an example of SAS code and output showing Michigan’s use of the Arthritis Management module. The SAS output shows that the frequencies of 1,511 in the second version and 1,634 in the third version of combined landline telephone and cellular telephone data are similar.

SAS code:

```
proc freq data=in.llcp13v2;
  where ARTHWGT ne . and _state=26;
  *ARTHWGT is one of the variables in the Arthritis Management Module';
  tables _state;
  title1 "Combined landline and cell phone multiple Version 2:";
  title2 "Michigan sample size for Arthritis Management Module";
run;
proc freq data=in.llcp13v3;
  where ARTHWGT ne . and _state=26;
  *ARTHWGT is one of the variables in the Arthritis Management Module';
  tables _state;
  title1 "Combined landline and cell phone multiple Version 3:";
  title2 "Michigan sample size for Arthritis Management Module";
run;
```

SAS Output:

Combined landline and cell phone multiple Version 2:  
Michigan sample size for Arthritis Management Module

The FREQ Procedure

STATE FIPS CODE				
_STATE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
26	1511	100.00	1511	100.00

Combined landline and cell phone multiple Version 3:  
Michigan sample size for Arthritis Management Module

The FREQ Procedure

STATE FIPS CODE				
_STATE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
26	1634	100.00	1634	100.00

## Example SAS code:

```
/*Example SAS Code*/
*Note: for this module, Michigan collected both Version 2 and Version 3 in the
combined landline
and cell phone;
*extract states from combined landline and cell phone data, common version;
data llcp; *observation number=69,940;
  *originally downloaded data from combined landline and cell phone data, common
version;
  set in.llcp2013 (where=(state in (21,27,28,29,42,47,49)));
  *state FIPS codes:
21(Kentucky),27(Minnesota),28(Mississippi),29(Missouri),42(Pennsylvania),
  47(Tennessee),49(Utah);
  *rename final weight variable to be consistent across new data sets;
  _finalwt=_llcpwt;
  drop _llcpwt;
run;

*extract states from combined landline and cell phone, Version 2;
data llcpv2; *observation number=21,579;
  *original downloaded data, select states that collected data from combined
landline and cell phone, Version 2;
  set in.llcp13v2 (where=(state in (6,20,30,36,41)));
  *state FIPS codes: 6(California),20(Kansas),30(Montana),36(New York),41(Oregon);
  *rename final weight variable to be consistent across new data sets;
  _finalwt = _lcpwtv2;
  drop _lcpwtv2;
run;

*extract Michigan because it collected this module from 2 multiple versions of
combined landline and cellphone data;
data MI2; *observation number=4,118;
  *original downloaded data, select Michigan that collected data from combined
landline and cell phone, Version 2;
  set in.llcp13v2 (where=(state=26));
  *rename final weight variable to be consistent across new data sets;
  _finalwt = _lcpwtv2/2;
  *Note: Version 2 weight is divided by 2 because Michigan collected 2 versions and
the sample size is same as Version 3 weight;
  drop _lcpwtv2;
run;
data MI3; *observariton number=4,335;
  *original downloaded data, select Michigan that collected data from combined
landline and cell phone, Version 3;
  set in.llcp13v3 (where=(state=26));
  *rename final weight variable to be consistent across new data sets;
  _finalwt = _lcpwtv3/2;
  *Note: Version 3 weight is divided by 2 because Michigan collected 2 versions and
the sample size is same as Version 2 weight;
  drop _lcpwtv3;
run;

*combine all data sets together;
data arthritis; *observation number=99,972;
  set llcp llcpv2 MI2 MI3;
run;
```