

Behavioral Risk Factor Surveillance System

Module Data for Analysis for 2012 BRFSS

July 15, 2013



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 persons residing in each of the states and selected U.S. territories. The BRFSS data collection, structure, and weighting methodology changed in 2011 to allow data to be collected by cellular telephones, in addition to landline telephones. 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. Analysis using BRFSS data should be conducted using complex sampling analyses. Data should be appropriately stratified and weighted in the analyses. Weighting can help selection probabilities and noncoverage among segments of the population.

If users intend to analyze the variables from the core section only, `_LLCPWT` is the appropriate weight for analysis. This document has been created to guide users who are analyzing the variables for 2012 BRFSS data using the module or module and core sections. Compared with 2011 data, split versions of module data collected on cellular telephones was introduced in 2012. Please note that the data set naming convention and weighting variables have changed from previous years in order to illustrate that weighting variables are different from those used in the past. Data users should become familiar with the information presented in this document prior to

performing analyses. More information about the changes to the 2012 BRFSS is available at http://www.cdc.gov/brfss/annual_data/2012/pdf/Compare_2012.pdf

Using BRFSS Data from Multiple Datasets

For 2012 BRFSS data, there are 8 datasets:

- 2012 BRFSS questionnaire data (combined landline telephone and cellular telephone).
- 2012 BRFSS combined landline telephone and cellular telephone multiple version questionnaire version 1 data (landline telephone and cellular telephone version 1).
- 2012 BRFSS combined landline telephone and cellular telephone multiple version questionnaire version 2 data (landline telephone and cellular telephone version 2).
- 2012 BRFSS combined landline telephone and cellular telephone multiple version questionnaire version 3 data (landline telephone and cellular telephone version 3).
- 2012 BRFSS landline telephone data (landline telephone only)
- 2012 BRFSS landline telephone multiple version questionnaire version 1 data (landline telephone version 1).
- 2012 BRFSS landline telephone multiple version questionnaire version 2 data (landline telephone version 2), and
- 2012 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 a state's needs. Although core questions are always collected on both landline telephone and cellular telephones, module data may have been collected by landline telephone and/or cellular telephone. In addition to 2011 data where states collected module data by landline telephone only, states also collected module data by combined landline telephone and cellular telephone in 2012. 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. The common modules have to be collected on all versions of combined landline telephone and cellular telephone or landline telephone-only data.

Prior to using data that have been collected in optional modules, users must identify the states that collected the data of interest and determine which questionnaire version the state used, if any. A listing of modules is available by state at <http://apps.nccd.cdc.gov/BRFSSModules/ModByState.asp?Yr=2012> and by module at <http://apps.nccd.cdc.gov/BRFSSModules/ModByCat.asp?Yr=2011>

As a first step to conduct 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. Keep in mind that there are eight possible data sets from which you will need to include data:

- 1) The combined landline telephone and cellular telephone data are used if the questions are exclusively from the core section or common modules asked on both the landline survey and cellular telephone survey.
- 2) The landline telephone survey data are used if the questions are asked only in the common version of the landline telephone survey.
- 3) 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, the variable `_STSTR` should be used for stratification and `_PSU` should be used for cluster in complex sampling analyses. The description of the data, name of the datasets, and the variable names of the final weight are illustrated in the following table.

Data description	Dataset name	Final weight variable name
Combined landline telephone and cellular telephone	LLCP2012	<code>_LLCPWT</code>
Combined landline telephone and cellular telephone Version 1	LLCP12V1	<code>_LCPWTV1</code>
Combined landline telephone and cellular telephone Version 2	LLCP12V2	<code>_LCPWTV2</code>
Combined landline telephone and cellular telephone Version 3	LLCP12V3	<code>_LCPWTV3</code>
Landline telephone only	LAND2012	<code>_LANDWT</code>
Landline telephone only Version 1	LAND12V1	<code>_LNDWTV1</code>
Landline telephone only Version 2	LAND12V2	<code>_LNDWTV2</code>
Landline telephone only Version 3	LAND12V3	<code>_LNDWTV3</code>

Step-by-Step Examples of Preparing 2012 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, always go through all documents on the relative to the given data year before conducting analysis and double-check state FIPS codes in each data set to avoid duplication.

Step-by-Step Example 1: Childhood Asthma Prevalence Module

The example below uses the Childhood Asthma Prevalence Module to demonstrate how to combine and reweight data from multiple data sets that can be used for analysis. This module was selected 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 2012 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 Childhood Asthma Prevalence module recorded in the combined landline telephone-cellular telephone dataset. These states are Alabama (1*), Arizona (4), California (6), Connecticut (9), District of Columbia (11), Georgia (13), Hawaii (15), Illinois (17), Kansas (20), Kentucky (21), Maryland (24), Michigan (26), Mississippi (28), Missouri (29), Montana (30), Nebraska (31), Nevada (32), New Hampshire (33), New Jersey (34), New Mexico (35), New York (36), Ohio (39), Oklahoma (40), Oregon (41), Pennsylvania (42), Puerto Rico (72), Tennessee (47), Texas (48), Utah (49), Washington (53), West Virginia (54), Wisconsin (55), Wyoming (56). (*Note: Numbers in parentheses are the State FIPS codes. The state FIPS codes can be found in <http://www.bls.gov/lau/lausfips.htm>.)

2. Search for states that collected data for this module with questionnaire version(s) in the 2012 landline telephone and cellular telephone data:

Go to the Web page for the [combined landline and telephone multiple version questionnaire data sets](#) and click on the document named “[Use of BRFSS Combined Land Line and Cellular Telephone Multiple Version Questionnaire Data](#).” This document lists modules by state and states by module. Go to page 4 where the states are listed under the module of Childhood Asthma Prevalence. Maine (23) is the state that collected data on Version 2 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-only questionnaire data sets](#) and click on the document named “[Use of BRFSS Landline Questionnaire Data](#)”; Go to the table listed by modules, which begins on page 4. Search for “Childhood Asthma Prevalence” 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 Louisiana (22), North Dakota (38), and Rhode Island (44) are the states that used this common module in the landline telephone only data.

4. Search for states that collected data for this module using multiple versions among landline telephone-only users:

Go to the Web page for [landline multiple version questionnaire data sets](#) and click on the document named “[Use of BRFSS Landline Multiple-Version Questionnaire Data](#)” and find the table of Landline Telephone Survey Questionnaire States and Modules. The table is listed by state and by module. Under the table listed by module, search for “Childhood Asthma Prevalence” to find the state that used multiple versions of the landline telephone- only survey—which is Massachusetts (25)—that uses questionnaire version 1.

5. Obtain data sets:

- To download the landline telephone and cellular telephone data set (LLCP2012):
Go to the [data Web page](#). From there, please read all the 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 2012 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 version 2:

Go to the [combined landline and cellular telephone multiple version questionnaire module data Web page](#) (add web link) 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: LLCP12V1.

- To download the landline telephone-only dataset (LAND2012):
[Go to the landline questionnaire data set Web page](#). The Web site also provides access to the 2012 landline telephone [multiple version questionnaires](#) or click on the Web site below, in the next bullet item.
 - To download multiple-version landline telephone data:
Go to the [data Web page](#) to download [multiple-version landline telephone data](#). There are 3 separate SAS data sets corresponding to version 1, version 2, and version 3. Choose Version 1: LAND12V1
6. Generate a new and uniform final weight variable from each of the data sets and combine data into a working data set. This will require users to:
- a. keep all the states that collect the module in each of the datasets,
 - 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 below.

Example SAS code:

```
*extract states from landline and cell phone combined data;
data llcp; *observation number=307,753;
  *originally downloaded data selecting states that collected data using Childhood
  Asthma Prevalence Module from combined landline and cell phone data for common
  module;
  set libname.llcp2012 (where=(state in
(1,4,6,9,11,13,15,17,20,21,24,26,28,29,30,31,
32,33,34,35,36,39,40,41,42,72,47,48,49,53,54,55,56)));
  *state FIPS code:
1(Alabama),4(Arizona),6(California),9(Connecticut),11(District of
Columbia),13(Gerogia),15(Hawaii),17(Illinois),20(Kansas),21(Kentucky),24(Maryland),
26(Michigan),28(Mississippi),29(Missouri),30(Montana),31(Nebraska),32(Nevada),
33(New Hampshire),34(New Jersey),35(New Mexico),36(New
York),39(Ohio),40(Oklahoma),41(Oregon),42(Pennsylvania),72(Puerto
Rico),47(Tennessee),48(Texas),49(Utah),53(Washington),54(West
Virginia),55(Wisconsin),56(Wyoming);
  _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;
data llcpv2; *observation number=4,962;
  *originally downloaded data selecting states that collected data for the
  Childhood Asthma Prevalence Module from combined landline and cell phone data,
  multiple versions;
  set libname.llcp12v2 (where=(state=23));
  *State FIPS code: 23(Maine);
  _finalwt=_lcpwtv2;
  *rename final weight variable to be consistent across new data sets;
  drop _lcpwtv2;
run;

*extract states from landline only data, common module;
data land; *observation number=16,431;
  *originally downloaded data, selecting state that collected data for the
  Childhood Asthma Prevalence Module from landline only data, common module;
  set libname.land2012 (where=(state in (22,38,44)));
  *State FIPS code: 22(Louisiana),38(North Dakota),44(Rhode Island);
  _finalwt = _landwt;
  *rename final weight variable to be consistent across new data sets;
  drop _landwt;
run;

*extract states from landline data, multiple Version 1 only;
data landv1; *observation number=6,077;
  *originally downloaded data selecting states that collected data for the
  Childhood Asthma Prevalence Module from landline only data, Questionnaire Version
  1;
  set libname.land12v1 (where=(state=25));
  *State FIPS code:25(Massachusetts);
  _finalwt = _lndwtv1;
  *rename final weight variable to be consistent across new data sets;
  drop _lndwtv1;
run;

*combine all data sets together;
```

```
data childasthma; *observation number=335,223;
  set llcp llcpv2 land landv1;
run;
```

Step-by-Step Example 2: Visual Impairment and Access to Eye Care Module

The example below uses the Visual Impairment and Access to Eye Care Module to demonstrate how to combine and reweight data from multiple datasets. 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. It appears that no states collected the data as a common module.
- 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 and cellular telephone multiple version questionnaire module data](#) Web page and click on the document titled, “[Use of BRFSS Combined Landline and Cellular Telephone Multiple Version Questionnaire Data](#)” and find table listed by module. Under “Visual Impairment and Access to Eye Care”, note that California (6) collected data for this module in both version 1 and version 2; Ohio (39) collected data for this module in version 1.
- 3. Search for states that collect this module for landline telephone only, common module:**
Go to the [Landline Questionnaire Data Sets](#) Web page and click on the document titled “[Use of BRFSS Landline Questionnaire Data](#)” and find table listed by module. There is no “Visual Impairment and Access to Eye Care” module, indicating this is not a common module in the landline telephone-only data.
- 4. Search for states that used this module for landline telephone-only participants, multiple versions:**
Go to the 2012 [Landline Multiple Version Questionnaire Data Sets](#) Web page and click on document titled, “[Use of BRFSS Landline Multiple Version Questionnaire Data](#)” and find the table listed by module. Under “Visual Impairment and Access to Eye Care,” note that Massachusetts (25) collected this module in version 3.
- 5. Obtain datasets:**
 - To download the landline telephone and cellular telephone dataset with multiple versions:
Go to the [Combined Landline and Cellular Telephone Multiple Version Questionnaire Module Data](#) Web page at (add Web link). 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 1 and version 2. The Web site also provides the link to access [2012 BRFSS landline multiple version telephone data](#) or click on the Web link in the next bulleted item, below.
 - To download multiple-version landline telephone data:**
Go to the 2012 [Landline Multiple Version Questionnaire Data Sets Web page](#). There are 3 separate SAS data sets corresponding to Version 1, Version 2, and Version 3. Download Version 3 data.
- 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.,

Visual Impairment and Access to Eye Care) where the states have multiple versions of combined landline telephone and cellular telephone data or landline telephone-only 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 California's use of the Visual Impairment and Access to Eye Care module. The SAS output shows that the frequencies of 2,817 in the first version and 2,728 in the second version of combined landline telephone and cellular telephone data are similar.

SAS code:

```
proc freq data=in.llcp12v1;
  where VIGLUMA2 ne . and _state=6;
  *VIGLUMA3 is one of the variables in the Visual Impairment and Access to Eye Care
Module';
  tables _state;
  title1 "Landline only Version 1:";
  title2 "California sample size for Visual Impairment and Access to Eye Care
Module";
run;
proc freq data=in.llcp12v2;
  where VIGLUMA2 ne . and _state=6;
  *VIGLUMA3 is one of the variables in the Visual Impairment and Access to Eye
Care Module';
  tables _state;
  title1 "Landline only Version 1:";
  title2 "California sample size for Visual Impairment and Access to Eye Care
Module";
run;
```

SAS Output:

Combined landline telephone and cellular telephone multiple Version 1:

California sample size for Visual Impairment and Access to Eye Care Module

The FREQ Procedure

STATE FIPS CODE

_STATE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
6	2817	100.00	2817	100.00

Combined landline telephone and cellular telephone multiple Version 2:

California sample size for Visual Impairment and Access to Eye Care Module

The FREQ Procedure

STATE FIPS CODE

_STATE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
6	2728	100.00	2728	100.00

Example SAS code:

```
*Note: for this module, there is no state collecting landline only or combined
landline and cell phone for common modules.
  Instead, California collected both Version 1 and Version 2 in the combined
landline and cell phone. Ohio collected data in
  the combined landline and cell phone Version 1. Massachusetts collected data in
the landline only data Version 3;
*extract Massachusetts from landline only, Version 3;
*extract Ohio from combined landline and cell phone, Version 1;
data llcpv1; *observation number=9,565;
  *originally downloaded data, select Ohio that collected data from combined
landline and cell phone, Version 1;
  set libname.llcp12v1 (where=(_state=39));
  *rename final weight variable to be consistent across new data sets;
  _finalwt = _lcpwtv1;
  drop _lcpwtv1;
run;

data landv3; *observation number=6,057;
  *originally downloaded data, select Massachusetts that collected data from
landline only Version 3;
  set libname.land12v3 (where=(_state=25));
  *rename final weight variable to be consistent across new data sets;
  _finalwt = _lndwtv3;
  drop _lndwtv3;
run;

*extract California because it collected this module from 2 multiple versions of
combined landline and cellphone data;
data CA1; *observation number=3,780;
  *originally downloaded data, select California that collected data from combined
landline and cell phone, Version 1;
  set libname.llcp12v1 (where=(_state=6));
  *rename final weight variable to be consistent across new data sets;
  _finalwt = _lcpwtv1/2;
  *Note: Version 1 weight is divided by 2 because California collected 2 versions
and the sample size is same as
  Version 2 weight;
  drop _lcpwtv1;
run;

data CA2; *observation number=3,635;
  *originally downloaded data, select California that collected data from combined
landline and cell phone, Version 2;
  set libname.llcp12v2 (where=(_state=6));
  *rename final weight variable to be consistent across new data sets;
  _finalwt = _lcpwtv2/2;
  *Note: Version 2 weight is divided by 2 because California collected 2 versions
and the sample size is same as
  Version 1 weight;
  drop _lcpwtv2;
run;

*combine all data sets together;
data Visual; *observation number=23,037;
  set llcpv1 landv3 CA1 CA2;
run;
```