Behavioral Risk Factor Surveillance System
Direct Age Adjustment, 2016 BRFSS
May 2017

Overview

The Behavioral Risk Factor Surveillance System (BRFSS) is a state-based, CDC-assisted health-data collection project and partnership of state health departments, CDC's Division of Population Health, and other CDC programs and offices. It comprises telephone surveys conducted by the health departments of all 50 states, the District of Columbia, Puerto Rico, the Virgin Islands and Guam. BRFSS collects 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. These updated procedures have been used each year afterward. The BRFSS includes a core set of questions and multiple optional modules that focus on specific health issues. Not all optional 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 and weight the data in their work. Potential bias resulting from selection probabilities and noncoverage among segments of the population can be reduced through weighting. Researchers conducting analysis of variables from the *core-only* section should use the variable **_LLCPWT** for weighting. The variable **_STSTR** should be used for stratification, and the variable **_PSU** should be used as a primary sampling unit.

The purpose of this document is to guide users who wish to conduct **direct age adjustment** (also called age standardization) using BRFSS data. Age adjustment is a standard analytical technique used to compare estimates between populations with different age distributions (e.g., between states) and over time. Additional details are provided on the following pages.

The direct age adjustment example in this document uses a variable from the core section of the questionnaire. The main BRFSS dataset, **LLCP2016**, contains information from the *core section* of the questionnaire. As described above, the variable **_LLCPWT** should be used for weighting this question from the core questionnaire. (Information on analyzing data from optional modules can be found in a separate document. Different weighting procedures may be required for optional module data; those details are beyond the scope of this document.)

The **LLCP2016** BRFSS data set includes data that respondents provided by landline telephone or 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, to 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 2016 BRFSS is available in the 2016 Data Comparability Report.

Prior to conducting analyses, users should always review all relevant documents for the given data year (e.g., the 2016 Overview Document, 2016 Codebook, etc.). Users should also double-check state FIPS codes in each data set to avoid duplication.

DIRECT AGE ADJUSTMENT: EXAMPLE

Background: Age adjustment is a standard analytical technique used to compare estimates between populations with different age distributions (e.g., between states) and over time. Direct age adjustment allows users to compare estimates across states and regions with different age distributions.
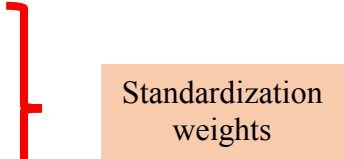
BRFSS calculates <u>direct</u> age-adjusted prevalence estimates using:

1. Age-specific prevalences in the <u>BRFSS data set</u>
2. Age distribution of an <u>external</u> standard population

BRFSS uses the **2000 Projected US Population** as the external standard population (1). This is consistent with the National Center for Health Statistics' current recommendation.

The age distribution of the 2000 Projected US <u>Adult</u> Population is as follows:

Age Category	Proportion of the 2000 Projected US Adult Population
18–44 years	0.530534557
45–64 years	0.299194019
≥65 years	0.170271424

Standardization weights

The 3 age categories and standardization weights above are used in the direct age adjustment of most BRFSS questions. Most BRFSS questions are asked of males and females aged ≥18 years.

Some BRFSS questions, however, are asked only of males, only of females, and/or only of certain age groups. In these instances, the categories and standardization weights used in direct

age adjustment are different. Selected examples are found below and represent only a couple of possibilities:

Age & Sex of BRFSS Respondents	Categories Used in Age Standardization	Proportion of 18 to 64 Year Old 2000 US Projected Population	Example BRFSS Variable
18–64 years, Both Male & Female	18–44 years	0.639407358	_HCVU651 (Health Care Coverage)
	45–64 years	0.360592642	

Please see next page for a detailed example.

Direct Age Adjustment Example: Part 1

Part 1 of this example will demonstrate how to:

- o Estimate the **age-adjusted** prevalence of health-care coverage among 18–64-year-old Minnesota residents in 2016 (variable: **_HCVU651**).

The steps to estimate this prevalence are listed below:

1. **Examine the variable coding in the 2016 Codebook and/or Calculated Variables document.**

Section 3 Calculated Variables

Variable Name	Description	Values
_HCVU651	Respondents aged 18–64 who have any form of health-care coverage	1 = Have health-care coverage 2 = Do not have health-care coverage 9 = Don't know/Not Sure, Refused or Missing, or age $\geq$ 65

Notes:
- _HCVU651 was calculated using variables from Section 3 of the **core** questionnaire.
- Adults aged ≥65 years are coded as _HCVY651= 9.
  - o This information can be found in the 2016 Calculated Variables document.

2. **Examine BRFSS age standardization categories and standardization weights.**
   - 63.9% of the 18–64-year-old 2000 US Projected Population was between 18 and 44 years (standardization weight for 18–44 year olds= 0.6394).
   - 36.1% of the 18–64-year-old 2000 US Projected Population was between 45 and 64 years (standardization weight for 45–64 year olds= 0.3606).

Age & Sex of BRFSS Respondents	Categories Used in Age Standardization	Proportion of 18–64-year-old 2000 US Projected Population (Standardization Weight)
18–64 years, Male/Female	18–44 years	0.639407358
	45–64 years	0.360592642

Direct Age Adjustment Example: Part 1 (continued)

3. **Write and execute code in SUDAAN (proc descript).**

```
libname us16 'BRFSS'; *Define where dataset is stored;

data mn16finl;
set us16.llcp2016 (where=(_state =27)); *Extract MN data from
landline and cell phone combined data (MN state FIPS code=27);


*Create a dichotomous variable representing the age
standardization categories (18-44 and 45-64). This dichotomous
variable will be derived from _age_g, a six-level variable with
the following values:
      1 (imputed age between 18-24 years)
      2 (imputed age between 25-34 years)
      3 (imputed age between 35-44 years)
      4 (imputed age between 45-54 years)
      5 (imputed age between 55-64 years)
      6 (imputed age between 65-99 years);

if 1 <= _age_g <=3 then age1864=1;      *Ages 18-44;
else if 4 <= _age_g <=5 then age1864=2; *Ages 45-64;
else age1864=.;                         *Ages 65 and above are
                                 coded as missing for the age
                                 standardization variable;

proc freq; tables _age_g*age1864/list missing; run;

proc sort data=mn16finl;     *Sort the data by the nest variables;
  by _ststr _psu; run;

proc descript data=mn16finl filetype=sas design=wr;
nest _ststr _psu /missunit;
weight _llcpwt;            *_HCVU651 was from core questionnaire.
                            Use standard analysis weight;
var _hcvu651;
catlevel 1;

subpopn _hcvu651 <9; *Among those with non-missing information;

  subgroup age1864;                  *Age standardization variable;
  levels   2;        *Age standardization variable has 2 levels;

  stdvar age1864;                    *Age standardization variable;
  stdwgt 0.639407358  0.360592642;      *Standardization weights;
```

Direct Age Adjustment Example: Part 1 (continued)

```
print nsum percent sepercent lowpct uppct/nsumfmt=F10.0
percentfmt=F20.4 sepercentfmt=F20.4 lowpctfmt=F20.4
uppctfmt=F20.4 style=NCHS;

title 'Age-adjusted prevalence of healthcare coverage among 18-64
year-olds, Minnesota, 2016';
run;
```

SUDAAN Output:

Age-adjusted prevalence of healthcare coverage among 18-64 year-olds, Minnesota, 2016

Sample size	Prevalence estimate (%)	Standard error (%)	Lower 95% CL (%)	Upper 95% CL (%)
11,637	92.0466	0.3307	91.3738	92.6711

CL, Confidence Limit.
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: _HCVU651 < 9
Standardized estimates

Note: This can be compared with the crude prevalence of 92.2671% (95% CI 91.6271%, 92.8620%); the age-adjusted prevalence is not very different from the crude prevalence in this case.

The code to calculate the crude prevalence is found below:

```
proc descript data=mn16finl filetype=sas design=wr;
nest _ststr _psu /missunit;
weight _llcpwt;
var _hcvu651;
catlevel 1;
subpopn _hcvu651 <9;

print nsum percent sepercent lowpct uppct/nsumfmt = F10.0
percentfmt=F20.4 sepercentfmt = F20.4 lowpctfmt=F20.4
uppctfmt=F20.4 style=NCHS;

title 'Crude prevalence of healthcare coverage among 18-64 year-
olds in Minnesota BRFSS, 2016';

run;
```

Please see next page for Part 2 of this example.

Direct Age Adjustment Example: Part 2

Part 2 of this example will demonstrate how to:

o Estimate the age-adjusted prevalence of health-care coverage among **18–64**-year-old Minnesota residents in 2016, **by race/ethnicity category** (_RACEGR3).

The steps to estimate this prevalence are listed below:

1. **Examine values of the _RACEGR3 variable.**

   - Notice that _RACEGR3=9 means "Don't know/Refused/Missing."
   - These individuals will be removed from the prevalence denominator.

Value of _RACEGR3	Race/Ethnicity
1	White (only), non-Hispanic
2	Black (only), non-Hispanic
3	Other race (only), non-Hispanic
4	Multiracial, non-Hispanic
5	Hispanic
9	Don't know/Refused/Missing

Notes:
   o The outcome variable (_HCVU651) is the same as in Part 1 of this example.
   o The age standardization categories and weights are the same as in Part 1 of this example.

2. **Write and execute code in SUDAAN (proc descript).**

```
proc descript data=mn16finl filetype=sas design=wr;
 nest _ststr _psu /missunit;
 weight _llcpwt;

 var _hcvu651;
 catlevel 1;

 subpopn _hcvu651 <9 and _racegr3 <9;  *Among those with non-
                                         missing information;

 subgroup age1864  _racegr3;  *Add _racegr3 to subgroup statement;
 levels      2        5;      *_Racegr3 has 5 (non-missing)
                                 categories;
```

```
tables _racegr3;        *Results by race/ethnicity category;

 stdvar age1864;        *Age standardization variable;
 stdwgt 0.639407358  0.360592642;      *Standardization weights;

 print nsum percent sepercent lowpct uppct/nsumfmt=F10.0
percentfmt=F20.4
sepercentfmt=F20.4 lowpctfmt=F20.4 uppctfmt=F20.4 style=NCHS;

 title 'Age-adjusted prevalence of healthcare coverage among 18-
64 year-olds by Race/Ethnicity, Minnesota, 2016';
 run;
```

SUDAAN Output:

Age-adjusted prevalence of health-care coverage among 18–64 year olds by Race/Ethnicity, Minnesota, 2016

Race/Ethnicity Category	Sample Size	Prevalence Estimate (%)	Standard Error	Lower 95% CL (%)	Upper 95% CL (%)
Total	11,501	92.0537	0.3327	91.3765	92.6819
White (only), non-Hispanic	9,847	94.8826	0.2956	94.2712	95.4319
Black (only), non-Hispanic	411	85.3042	2.0112	80.9104	88.8263
Other race (only), non-Hispanic	513	90.1126	1.6675	86.3299	92.9343
Multiracial, non-Hispanic	134	90.8803	2.6014	84.3416	94.8551
Hispanic	596	60.7555	2.3742	56.0174	65.2994

CL: Confidence Limit.
Variance Estimation Method: Taylor Series (WR)
For Subpopulation: _HCVU651 < 9 AND _RACEGR3 <9
Standardized estimates

Reference
Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People 2010 Stat Notes. 2001(20):1-10.