

## **Technical Notes**

---

### **Survey Management**

The Behavioral Risk Factor Surveillance System is a cooperative effort of the Centers for Disease Control and Prevention and participating states. The CDC develops the core questionnaire and provides training, technical assistance, standardized data analyses and funding. The Pennsylvania Department of Health develops supplemental questions (or modules) and responds to requests for data. In 2010, Pennsylvania was divided into eight strata. These strata consisted of eight regional areas: six Pennsylvania health districts and Allegheny and Philadelphia Counties. Sampling and interviewing in 2010 were done by Clearwater Research, Incorporated, which was selected by competitive bid.

### **Sample Selection**

Respondents were selected using a two-stage random digit dialing sample design. In the first selection stage, a disproportionate stratified sample of telephone numbers was selected from two telephone number strata. One stratum consists of listed Pennsylvania residential telephone numbers. The other stratum consists of blocks of telephone numbers, known to include one or more residential telephone numbers. Both strata's telephone numbers begin with area codes and exchange prefixes specific to Pennsylvania. Cell phone numbers are not included in the 2010 BRFSS sample selection.

Under the disproportionate stratified sample design, the sample of Pennsylvania telephone numbers that is divided into the two strata described in the previous paragraph are selected based on the estimated probability that the phone number is attached to a housing unit. A larger proportion of the sample is selected from the strata of listed telephone numbers known to be connected to residential households.

In the first stage of sampling, the selected telephone numbers were called to determine if they were residential telephone numbers. Nonresidential telephone numbers were discarded from the sample. Residential numbers were subjected to the second stage of sampling, wherein an adult was randomly selected as the respondent from a list of adults residing in the household. The person who answers the telephone generates this list.

### **Questionnaire**

The survey questionnaire for the statewide survey of Pennsylvania consists of a standardized core, state-selected modules and state-added questions. The CDC developed the core questionnaire with recommendations from all participating jurisdictions. Most of the core questions had been used during the 2009 BRFSS survey. All items new to the 2010 survey were field-tested.

Questions of interest to Pennsylvania were added as the state supplement to the core questionnaire. State-added modules and questions concerned diabetes, asthma, general preparedness, childhood immunization, adverse childhood experience, Child Human Papilloma Virus (HPV), gambling, worksite wellness, dental sealants, colorectal screening and doctor advice on smoking.

### **Response Rates**

Interviews were conducted in the evenings and on weekends in order to reach people when they were more likely to be at home, as well as during the day. At least 15 calls were placed at different times of the day and night during different days of the week before any sample number was classified as "no answer."

Interviewers who were experienced in converting refusals to completed interviews re-contacted people who refused to participate in the survey.

The outcome of all telephone calls is shown on the following page. This report was prepared using data from 10,543 completed interviews. The CASRO response rate for 2010 was 46 percent.

**Disposition of All Telephone Numbers  
2010 Pennsylvania Behavioral Risk Factor Survey Sample**

<u>Disposition Code</u>	<u>Number</u>	<u>Percent</u>
<b>Interview</b>		
Complete	10,543	7.91
Partial Complete	673	0.51
<b>Eligible, Non-Interview</b>		
Termination within questionnaire	1,370	1.03
Refusal after respondent selection	3,680	2.76
Selected respondent never reached or was reached but did not begin interview during interviewing period	1,251	0.94
Selected respondent away from residence during the entire interviewing period	922	0.69
Language problem after respondent selection	144	0.11
Selected respondent physically or mentally unable to complete an interview during the entire interviewing period	637	0.48
Hang-up or termination after number of adults recorded but before respondent selection	47	0.04
<b>Unknown Eligibility, Non-Interview</b>		
Household members away from residence during entire interviewing period	244	0.18
Hang-up or termination, housing unit, unknown if eligible respondent	392	0.29
Household contact, eligibility undetermined	106	0.08
Language problem before respondent selection	401	0.30
Physical or mental impairment before respondent selection	291	0.22
Hang-up or termination, unknown if private residence	13,629	10.23
Contact, unknown if private residence	1,882	1.41
Telephone answering device, message confirms private residential status	1,521	1.14
Telecommunication technological barrier, message confirms private residential status	52	0.04
Telephone answering device, not sure if private residence	5,826	4.37
Telecommunication technological barrier, not sure if private residence	93	0.07
Telephone number is no longer in service or has been changed	1,132	0.85
No answer	4,634	3.48
Busy	318	0.24
<b>Not Eligible</b>		
Out-of-state	27	0.02
Household, no eligible respondent	15	0.01
Not a private residence	4,346	3.26
Dedicated fax/data/modem line with no human contact	3,644	2.73
Cellular Phone	329	0.25
Fast busy	1,534	1.15
Non-working/disconnected number	73,576	55.21
Total	133,259	

## Sample Characteristics

The following table compares the final interview sample for the 2010 BRFSS to the 2009 Population estimates for the adult population of Pennsylvania. The observations used to calculate the estimates presented in the main report were weighted to account for differences between the population and the distribution of age, sex, race and Hispanic origin characteristics in the sample.

**Distribution of 2010 Pennsylvania BRFSS Survey Sample and  
2009 Pennsylvania Adult Population Estimates For Selected Characteristics**

		2010 BRFSS Survey Sample		2009 Population Estimates	
		Number	Percent	Number	Percent
All Adults		11,211	100.00	9,755,408	100.00
Sex	Male	4,234	37.77	4,673,949	47.91
	Female	6,977	62.23	5,081,459	52.09
Race	White	9,939	88.65	8,461,285	86.73
	Black	864	7.73	974,584	9.99
	Other	184	1.64	319,539	3.28
	Unknown/Refused	221	1.97	NA	NA
Hispanic Origin	Yes	191	1.70	402,603	4.13
	No	10,960	97.53	9,352,805	95.87
	Unknown/Refused	86	0.77	NA	NA
Age	18-24	265	2.36	1,300,540	13.33
	25-34	793	7.07	1,507,763	15.46
	35-44	1,479	13.19	1,665,407	17.07
	45-54	2,050	18.29	1,813,742	18.59
	55-64	2,542	22.67	1,521,664	15.60
	65-74	2,007	17.90	977,199	10.02
	75+	1,966	17.54	969,093	9.93
	Unknown/Refused	109	0.97	NA	NA

**Note<sup>1</sup>:** Race data include Hispanics.

**Note<sup>2</sup>:** Population estimates allocate unknowns, so they are included in demographic categories. This is further indicated by the use of “NA” or not applicable for the 2009 population estimate “Unknown/Refused” entries.

### Determining Accuracy of the Estimates and Significance Using Confidence Intervals

Tables included in this report show the 95 percent confidence intervals associated with all reported percentages. They appear in the table columns labeled (CI).

Confidence intervals are a way to measure sampling error and define the range of values where percentages estimated by multiple samples of the same population would be found (95 percent of the time). The size of the confidence interval is directly related to the probability of selection and characteristics of the people surveyed within the universe being sampled. Percentages for two different subgroups of the population are significantly different if their confidence intervals or ranges do not overlap.

Confidence intervals were calculated using SUDAAN, a software package developed by the Research Triangle Institute, that properly estimates sample variances for complex sample designs.

Percentages were not calculated and shown for subgroups of the population when their sample size was less than 50. The method used to determine the reliability of percentages calculated from sample sizes of 50 or more consisted of a comparison of the relative standard error of the calculated percentage with the relative standard error of the same percentage outcome for a simple random sample. If the relative standard error for the percentage being tested was smaller than the relative standard error of the same percentage outcome for the simple random sample, then the calculated percentage was considered reliable.

### **Data Adjustment**

The data were adjusted, or weighted, to correspond to estimated age and sex distribution of the population for Pennsylvania in 2009. Because people living in households with more than one telephone or more than one adult had differing probabilities of being selected, the responses were also adjusted to reflect the number of different telephone numbers per household and the number of adults residing in the household. All of the percentages reported here were calculated with weighted data and should be representative of the adult population of Pennsylvania. It should be noted that the percentages might not add to 100, due to rounding. When calculating the percentages of prevalence for each health topic in this report, responses of “Don’t know/Not sure” and “Refused” were removed from the denominators.

### **Data Comparison to Other Sources**

Please note that different sources of behavioral risk factor surveillance system data for Pennsylvania may display percentages that are different from those displayed in this report. The Department’s interactive web tool Epidemiological Querying and Mapping System (EpiQMS) uses data from the Pennsylvania Behavioral Risk Factor Surveillance System. This report only uses the statewide sample to compute percentages. Also, other sources may include or exclude data from the denominator data used to calculate percentages. For example, the exclusion or inclusion of “Don’t Know/Not Sure” or “Refused” responses could potentially affect the final response percentage calculation.

If you have any questions about these differences, please contact the Bureau of Health Statistics and Research by phone at 717-783-2548 or by mail at 555 Walnut St., 6th Floor, Harrisburg, PA 17101-1914.

## Synthetic Estimation Process for Local Data

The BRFSS is an ongoing telephone survey consisting of interviews conducted each month. In 2009, the sample dataset includes 11,237 surveys divided into eight different Pennsylvania health regions (Allegheny and Philadelphia are separate).

On the state level, data from the BRFSS serve several purposes. BRFSS data help to identify subgroups, which should be targeted for health promotion and disease prevention programs due to elevated risks. Multiple years of BRFSS data are useful for tracking Pennsylvania's progress in achieving selected Healthy People 2010 National Health Objectives. Data from Pennsylvania, when compared to similar data from other states, identifies the need for increased health promotion and disease prevention program efforts. In 2010, comparable data were available from all 50 states, the District of Columbia, Puerto Rico, the U.S. Virgin Islands and Guam.

On the local level, BRFSS data may also be used to estimate the prevalence of risks in local areas such as counties, if the data are combined for several years or the counties or county groups of interest are over-sampled. However, for most counties, the number of respondents in the BRFSS sample data set is insufficient to produce reliable estimates.

In cases where local data on behavioral risk are not available, synthetic estimates can be computed based on either national data or statewide data from the BRFSS. Synthetic estimates are calculated using population estimates for subgroups of interest and the state or national risk factor prevalence rates for those groups. Below is an example of how one can compute synthetic estimates for a local area:

### Step 1

Obtain the population estimates for the local geographic area of interest. Sum the population estimates into a table with the same breakdown as a table listing the national or state estimates (see the table below).

### Step 2

To estimate the number of persons who have the behavioral risk in each subgroup, multiply the subgroup-specific rates by the population estimates for each group. For example, multiply the 2010 (latest available) Dauphin County census population of 40,985 for ages 18-29 by the 2010 fair or poor health prevalence of 9 percent (0.09) for that age group at the state level. The 2010 synthetic estimate for those in fair or poor health ages 18-29 in Dauphin County is 3,689.

### Step 3

To obtain the total number of persons who indicated fair or poor health, repeat Step 2 for all subgroups and then sum the subgroup estimates to get a total estimate.

Age Group	2007 Dauphin County Census Population		Fair or Poor Health from 2008 Pa. BRFSS		Estimate of Dauphin County Adults Indicating Fair or Poor Health, 2008
18-29	40,985	x	9 %	=	3,689
30-44	51,044	x	10 %	=	5,104
45-64	77,015	x	18 %	=	13,863
65+	36,841	x	26 %	=	9,579
				Total	32,235

**Caution:** Synthetic estimates can be useful for planning purposes. However, these estimates should not be used if there is reason to believe that local rates for subgroups of interest would diverge widely from the state or national rates. The prevalence of most health-related conditions varies considerably with age and often with other factors, such as sex, race and income. A more precise estimate may be obtained using age, sex and race-specific prevalence rates. The BRFSS is not a reliable source of prevalence rates specific to age-sex-race categories; national data would be a more reliable basis for synthetic estimates.

It is important to qualify estimates whenever they are used. A clear citation of the sources of the data used to compute the local area synthetic estimates should be included in every report of the estimates.

**Step 4**

To calculate the synthetic estimated percentage of Dauphin County adults with fair or poor health, pull the “Total Estimated Number of Adults” and the “Total Population Age 18+” in Dauphin County from “Step 3.”

Total Synthetically Estimated Number of Adults With Fair or Poor Health in Dauphin County = **32,235**

Total Population Age 18+ in Dauphin County = **205,885**

Divide the synthetically estimated number of adults with fair or poor health by the adult population. Then multiply by 100 so that the result will be expressed as a percent.

$$\text{Synthetically Estimated Percentage With Fair or Poor Health in Dauphin County} = \frac{\text{Total Synthetically Estimated Number of Adults With Fair or Poor Health in Dauphin County}}{\text{Total Population Age 18+ in Dauphin County}} \times 100$$

Synthetically Estimated Percentage With Fair or Poor Health in Dauphin County = (32,235 / 205,885) X 100

Synthetically Estimated Percentage With Fair or Poor Health in Dauphin County = **15.7 Percent**

This step gives you a synthetically estimated percentage of adults.

**Caution:** Synthetic estimates can be useful for planning purposes. However, these estimates should not be used if there is reason to believe that local rates for subgroups of interest would diverge widely from the state or national rates. The prevalence of most health-related conditions varies considerably with age and often with other factors, such as sex, race and income. A more precise estimate may be obtained using age, sex and race-specific prevalence rates. The BRFSS is not a reliable source of prevalence rates specific to age-sex-race categories; national data would be a more reliable basis for synthetic estimates.

It is important to qualify estimates whenever they are used. A clear citation of the sources of the data used to compute the local-area synthetic estimates should be included in every report of the estimates.

## References

---

Bureau of Health Statistics and Research; Behavioral Health Risks of Pennsylvania Adults 2009 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics and Research; Behavioral Health Risks of Pennsylvania Adults 2008 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics and Research; Behavioral Health Risks of Pennsylvania Adults 2007 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics and Research; Behavioral Health Risks of Pennsylvania Adults 2006 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics and Research; Behavioral Health Risks of Pennsylvania Adults 2005 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics and Research; Behavioral Health Risks of Pennsylvania Adults 2004 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics and Research; Behavioral Health Risks of Pennsylvania Adults 2003 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics and Research; Behavioral Health Risks of Pennsylvania Adults 2002 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics and Research; Behavioral Health Risks of Pennsylvania Adults 2001 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics; Behavioral Health Risks of Pennsylvania Adults 2000 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Bureau of Health Statistics; Behavioral Health Risks of Pennsylvania Adults 1999 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

Division of Health Statistics; Behavioral Health Risks of Pennsylvania Adults 1998 Harrisburg, Pa., Pennsylvania Department of Health; #303.300P.

U.S. Department of Health and Human Services. Healthy People 2010: Understanding and Improving Health. 2nd ed. Washington, DC: U.S. Government Printing Office, November 2000.

Van Belle, Gerald (2002). *Statistical Rules of Thumb*. Canada: Wiley.