Changes in the BRFSS Weighting Methodology

The Behavioral Risk Factor Surveillance System (BRFSS) is a cross-sectional telephone survey conducted by state health departments with technical and methodological assistance provided by the Centers for Disease Control and Prevention (CDC). The BRFSS survey consists of telephone interviews using randomly generated telephone numbers to determine the households contacted. The questions asked concern the leading causes of death, heart disease, cancer, diabetes, injury and other health-related behaviors.

Telephone surveys are weighted to reduce bias caused by non-coverage and non-response. Up until 2011, the data from the BRFSS survey were weighted by a standard method called post-stratification. Responses to the survey were forced to equal the age, sex and race proportions of adults in the household population. To be representative of the adult population of Pennsylvania, all the estimates reported are calculated with weighted data. The formula used to post-stratify BRFSS data before 2011 is:

FINALWT=STRWT*1/IMPNPH*NUMADULT*POSTSTR

Where FINALWT is the weight assigned to each respondent, STRWT is the inverse of the sampling fraction of each stratum and accounts for differences in the basic probability of selection among strata, IMPNPH is the number of residential telephone numbers in the respondent's house, NUMADULT is the number of adult's in the respondent's household and POSTSTR is the number of people in an age-by-sex or age-by-race-by-sex category in the population of a region or a state divided by the sum of the products of the preceding weights for the respondents in that same age-by-sex or age-by-race-by-sex category.

Post-stratification offered a means of providing the best possible prevalence estimates before 2011. However, societal changes often necessitate changes in survey methodology. Advancements in technology have had both positive and negative effects on the ability of telephone surveys, such as the BRFSS, to provide valid and representative data. The advent of extremely fast microprocessors for desktop computers and local area networks has enabled the routine use of more complex statistical weighting procedures to account for differences between survey respondents and the target population. The continuing trend towards the replacement of household landline telephones with personal cell phones in the United States has made it necessary to introduce new weighting measures.

The 2011 survey marks the first year in which the Pennsylvania BRFSS collected data from both landline and cell phone respondents. To allow for the incorporation of cell phone data, a new weighting methodology called iterative proportional fitting, or raking was implemented in 2011.

Raking is accomplished by adjusting for one demographic variable (or margin) at a time. For example, when weighting by age and gender, weights would first be adjusted for gender groups, then those estimates would be adjusted by age groups. This iterative process would continue until all group proportions in the sample approach those of the population, or up to 75 iterations. Raking allows for the incorporation of cell phone survey data, permits the introduction of additional demographic characteristics and more accurately matches sample distributions to known demographic characteristics of populations.

Rather than adjust the survey population to match the overall population on only three demographic variables, raking weights will use the following eight demographic dimensions, some of which are the intersection of two demographic subgroups:

- age group by gender
- detailed race/ethnicity

- educational level
- marital status
- home owner or renter status
- gender by race/ethnicity
- age group by race/ethnicity
- telephone source (landline telephone only, both landline and cell phone, or cell phone only)

For those states that use regional weighting, such as Pennsylvania's eight health-planning regions, the raking procedures include additional raking dimensions:

- region
- region by age group
- region by gender
- region by race/ethnicity

A principle advantage of raking is that it allows the use of many more adjustor variables than post-stratification. Together, the addition of the cell phone sample and the use of raking yield a more accurate weighting process and more precise estimates of the prevalence of health-related behaviors. The trend toward cell phone-only households is especially strong in younger age groups and among persons who are in racial and ethnic minority groups. The inclusion of the telephone source (landline or cell phone) in the weighting methodology allows the BRFSS to better represent lower-income and minority populations, factor in younger age groups and also better represent the lower levels of formal education within a population.

Due to the inclusion of cell phones and the change in methodology, 2011 and future BRFSS estimates should not be used to compare to estimates from 2010 and earlier and 2011-2013 and future state and regional estimates should not be used to compare to estimates from 2010-2012 and earlier. The 2011 estimates should be considered the new BRFSS baseline. These methodological changes will cause breaks in BRFSS trends but they will significantly improve the accuracy, coverage, validity, and representativeness of the Pennsylvania BRFSS.