COMPARABILITY OF DATA: BRFSS 2000

The BRFSS is a cross-sectional surveillance survey currently involving 52 reporting areas. The Virgin Islands and Guam are point-in-time surveys and are not discussed here. It is important to note that any survey will have natural variation over sample sites; therefore some variation between states is to be expected. The complex sample design and the multiple reporting areas complicate the analysis of the BRFSS. Although CDC works with the states to minimize deviations, in 2000 there were some deviations in sampling and weighting protocols, sample size, response rates, and collection or processing procedures. In addition, California=s questionnaire had a few minor differences in wording of question. The following section identifies other known variations for the 2000 data year.

A. 2000 Data Anomalies and Deviations from Sampling Frame and Weighting Protocols

In several states, a portion of sample records intended for use during one month were completed in another month. This deviation will disproportionately affect analyses based on monthly, rather than annual data.

Several states did not collect data for all twelve months of the year. Ohio did not collect data January through March, or July through September. Their data include interviews completed in January and February of 2001. Arizona is missing data in July, August, September and December. The District of Columbia did not start the survey until April.

Hawaii did not release telephone numbers in accordance with protocol and may have introduced bias into the survey. More information about the quality of the survey data can be found in the *2000 BRFSS Summary Data Quality Report*.

B. Other 2000 limitations of the data

Telephone coverage varies by state and also by subpopulation. Telephone coverage averages about 95% for U.S. states as a whole, but ranges from 1.8% non-coverage in Delaware, to 13.3% in New Mexico. It is estimated 10% of households in Puerto Rico are without telephones.

Dual questionnaires and/or partial year coverage occurred in Illinois. Illinois used a dual questionnaire and collected data on core items involving immunization, cholesterol, hypertension, colorectal screening, injury and alcohol consumption and modules concerning exercise, cigar smoking, fruits and vegetables and weight control for only six months of the interviewing period.

California modified the wording of mammography, clinical breast exam, and Papanicolaou (PAP) smear questions. More than 7%, (n=309) of Puerto Rico=s respondents did not answer the fruit and vegetable questions. These questions may have limited comparability to those of other reporting areas.

More information about survey item non-response can be found in the *2000 BRFSS Summary Data Quality Report*.

STATISTICAL AND ANALYTIC ISSUES

Estimation Procedures

Unweighted data on the BRFSS are the actual responses of each respondent. Unweighted data represent results before any adjustment is made for variation in respondents= probability of selection, for disproportionate selection of population subgroups relative to the state=s population distribution, or nonresponse. Weighted BRFSS data represent results that have been adjusted to compensate for such differences. Irrespective of state sample design, use of the final weight in analysis is necessary if generalizations are to be made from the sample to the population.

Statistical Issues

The procedures for estimating variances given in most statistical texts and the programs available in most statistical software packages are based on the assumption of simple random sampling (SRS). The data collected in the BRFSS are obtained through a complex sample design; therefore, the direct application of standard statistical analysis methods for variance estimation and hypothesis testing may yield misleading results. There are computer programs available which take such complex sample designs into account.   SUDAAN and EpiInfo=s C-Sample are among those suitable for analyzing BRFSS data. EpiInfo=s C-sample can be used to calculate simple frequencies and two-way cross-tabulations. SUDAAN can be used for both descriptive and inferential statistical methods. These software products require knowing the stratum, the primary sampling units, and the record weight; all of which are on the master data file. For more information on calculating variance estimation using SUDAAN, see Shah,1990 (1). For information about EpiInfo, see Dean, et al, 1995 (2).

Although the overall number of persons in the BRFSS is quite large for statistical inference purposes, subgroup analyses can lead to estimators that are unreliable. Consequently, analysis of subgroups, especially within a single data year or geographic area, requires that the user pay particular attention to the subgroup sample size. Small sample sizes may produce unstable estimates. Reliability of an estimate depends on the actual **unweighted** number of respondents in a category, not on the weighted number. Interpreting and reporting weighted numbers that are based on a small, unweighted number of respondents can mislead the reader into believing that a given finding is much more precise than it actually is. The BRFSS follows a rule of not reporting or interpreting percentages based upon a denominator of less than 50 respondents (unweighted sample).

The variable that tells SUDAAN about stratification in the BRFSS survey, **_STSTR**, does not identify density strata. The reason for the omission is that when density strata were identified, several strata contained too few records for SUDAAN to run properly. We checked on the effects of omitting density strata by comparing results with and without density strata for states with enough records for SUDAAN to run properly. The inclusion of density strata information made very little difference. Thus, it was decided to exclude this information from _STSTR in order to allow SUDAAN to run without unduly complicating the DESIGN statement.

Analytic Issues

Advantages and Disadvantages of Telephone Surveys

Compared with in-person interviewing techniques, telephone interviews are easy to conduct and monitor, and cost efficient. However, telephone interviews have limitations. Telephone surveys may have higher levels of noncoverage than in-person interviews because a percentage of U.S. households cannot be reached by telephone. As mentioned earlier, approximately 98 percent of households in the U.S. have telephones. A number of studies have shown that the telephone and nontelephone populations are different with respect to demographic, economic, and health characteristics (3,4,5). Although the estimates of characteristics for the total population are unlikely to be substantially affected by the omission of the nontelephone households, some of the subpopulation estimates could be biased due to the noncoverage of households without telephones. Telephone coverage is lower for population subgroups such as blacks in the South, persons with low incomes, persons in rural areas, persons with less than 12 years education, persons in poor health, and heads of households under 25 years of age (6). However, post-stratification adjustments for age, race, and sex, and other weighting adjustments used for the BRFSS data minimize the impact of differences in noncoverage, undercoverage, and nonresponse at the State level. State-specific information on telephone coverage is available in the technical documentation section on the BRFSS website, *www.cdc.gov/nccdphp/brfss* **.**

Despite the above limitations, prevalence estimates from the BRFSS correspond well with findings from surveys based on in-person interviews, including studies conducted by the National Institute on Alcohol Abuse and Alcoholism, CDC=s National Center for Health Statistics, and the American Heart Association (7). A summary of methodologic studies of BRFSS is provided in the publication section on the BRFSS website, *www.cdc.gov/nccdphp/brfss* .

Surveys based on self-reported information may be less accurate than those based on physical measurements. For example, respondents are known to underreport weight. Although this type of potential bias is an element of both telephone and in-person interviews, it should be considered by the analyst interpreting self-reported data. When measuring change over time, this type of bias is likely to be constant, and therefore not a factor in trend analysis.

Aggregating Data Over Time

When data from one time period are insufficient for estimating the prevalence of a risk factor, data may be combined for several periods as long as the periods being combined are not times during which the prevalence of the risk factor of interest has been substantially changing. One method that can be used to assess the stability of the prevalence estimates is discussed below (8).

1. Compute the prevalence for the risk factor for each period.

2. Rank the estimates from low to high.

3. Identify a statistical test appropriate for comparing the lowest and the highest estimates at the 5% level of significance. For example, depending on the type of data, a t-test or the sign test might be appropriate.

4. Test the hypothesis that prevalence is not changing by using a two-sided test in which the null hypothesis is that the prevalences are equal.

5. Determine whether the resulting difference could be expected to occur by chance alone less than 5% of the time (i.e., test at the 95% confidence level).

Analyzing Subgroups

When the prevalence of risk factors does not change rapidly over time, data combined for two or more years may provide a sufficient number of respondents so that additional prevalence estimates can be made for population groups (such as age/sex/race subgroups or county populations). Before combining data for subgroups, determine whether the total number of respondents will yield the precision needed. The level of precision needed depends upon the intended use of the estimate. For example, greater precision would be required to justify implementing expensive programs than that for general information only.

The table below shows the sample size required for each of several levels of precision based on a calculation in which the estimated risk factor prevalence is 50% and the design effect is 1.5.

Precision Desired	Sample Size Needed
2%	3600
4%	900
6%	400
8%	225
10%	144
15%	64
20%	36

Precision is indicated by the width of the 95% percent confidence interval around the prevalence estimate. For example, a desired precision of 2% means that the 95% confidence interval is + or - 2% of 50%, or 48-52%. As shown in the table, to yield this high a level of precision, the sample size required is about 3,600 persons. When a lower level of precision is acceptable, the sample size can be considerably smaller.

The design effect is a measure of the complexity of the sampling design and indicates how the design differs from simple random sampling. It is defined as the variance for the actual sampling design divided by the variance for a simple random sample of the same size (9, 10). For most risk factors in most states, the design effect is less than 1.5. If it is more than 1.5, however, sample sizes may need to be larger than those shown here.

The standard error of a percentage is largest at 50% and decreases as a percentage approaches 0% or 100%. From this perspective, the required sample sizes above are conservative estimates. They should be reasonably valid for percentages between 20% and 80% but may significantly overstate the required sample sizes for smaller or larger percentages.

As a cautionary note, users should remember that the reliability of an estimate depends on the actual, unweighted number of respondents in a category, not on the weighted number. Interpreting and reporting weighted numbers that are based on a small, unweighted number of respondents can mislead the reader into believing that a given finding is much more precise than it actually is. **The BRFSS strongly urges all users to follow the general rule of not reporting or interpreting percentages based upon a denominator with fewer than 50 unweighted respondents.**

Creating Synthetic Estimates

Sample sizes may still be inadequate for risk factor estimates for some geographic areas (i.e. counties) or subpopulations (i.e., persons with diabetes) even after combining data for several years. In such situations, the analyst may wish to derive synthetic estimates by extrapolating from the BRFSS data collected at the state level.

Synthetic estimates can be calculated by using the population estimates for the subgroup of interest and the state BRFSS risk factor prevalences for that subgroup. This approach assumes that the risk factor prevalences for specific subgroups in each area are the same as the statewide risk factor prevalences for the same subgroups. For example, it assumes that the risk factor prevalences for black women in every county of a state are the same as those for black women in the entire state. The accuracy of the estimate depends on the validity of this assumption, which is often impossible to judge. However, a Aballpark@ estimate may be sufficient for establishing broad goals and objectives for prevention strategies. For a discussion of the precision of such estimates, see Levy and Lemeshow (11).

An example for estimating the number of persons with hypertension in a hypothetical county, as well as the overall prevalence of hypertension in that county is shown below. The sex and race distribution of the county=s population differs from the statewide population, and these differences need to be taken into account. By developing a table like the one below, a synthetic estimate for the overall county prevalence of hypertension can be made.

Synthetic Estimates of Prevalence of Hypertension in a Hypothetical County, 1990

State Subgroup	Prevalence (per 100 Persons)	County Population	County Population with Hypertension
Men			
White	15.6	10,000	1,560
Black	27.0	25,000	6,750
Women			
White	19.5	12,000	1,340
Black	25.6	28,000	7,420
TOTAL		75,000	18,070

The state prevalence values, given as rates per 100 persons, are computed from the BRFSS data. The estimated number of persons with hypertension for each race-sex group in the county was obtained by multiplying the statewide prevalence for that group by the county population for the group. To determine the total county prevalence, the number of people with hypertension in each race-sex group in the county were summed and this sum (18,070) was divided by the county=s total population (75,000) to yield a overall prevalence of 24.1 per 100 persons.

Creating Direct Estimates

If the subpopulation sample size is sufficient to do so, analysts may choose to produce direct estimates. SUDAAN or a similar program will be needed for direct estimates. The subarea (i.e., county) is treated as a population domain for which the risk estimate is sought, and will be defined as a SUBGROUP variable in SUDAAN. Temporal and spatial stratification must be incorporated into the estimates of variable, by inclusion in the NEST statement in SUDAAN. If possible, it is desirable to re-adjust the poststratification weight (_POSTSTR) to the age-by-race-by gender population distribution of the small area (i.e., county).

To locally post-stratify the CDC BRFSS weights used for the direct estimate, post-stratify _WT1 to the population of interest. The equivalent local final weight is a product of _WT1 and the local poststratification factor.

REFERENCES

1. Shah BV. Software for Survey Data Analysis (SUDAAN), Version 6.1, Research Triangle Institute, Research Triangle Park, N.C. 1990.

2. Dean AG, Dean JA, Coulombier D, Brendel KA, Smith DC, Burton AH, Dicker RC, Sullivan K, Fagan RF, Arner TG. Epi Info, Version 6.0: A word processing, database, and statistics program for public health on IBM-compatible microcomputers. Centers for Disease Control and Prevention. 1995.

3. Groves RM, Kahn RL. Surveys by Telephone: A national comparison with personal interviews, New York, Academic Press, 1979.

4. Banks MJ. Comparing health and medical care estimates of the phone and nonphone populations. Proceedings of the Section on Survey Research Methods, American Statistical Association, 1983, pp 569-574.

5. Thornberry OT, Massey JT. Trends in United States Telephone Coverage Across Time and Subgroups. In Groves, RM et al editors Telephone Survey Methodology, pp. 25-49, New York, John Wiley & Sons, 1988.

6. Massey JT, Botman SL. Weighting Adjustments for Random Digit Dialed Surveys. In Groves, RM et al editors Telephone Survey Methodology, pp. 143-160,New York, John Wiley & Sons, 1988.

7. Frazier EL, Franks AL, Sanderson LM. Behavioral Risk Factor Data. In Using chronic disease data: A handbook for public health practitioners, pp 4.1-1.17. Centers for Disease Control and Prevention. 1992.

8. Frazier EL, Franks AL, Sanderson LM. Behavioral Risk Factor Data. In Using chronic disease data: A handbook for public health practitioners, pp 4.1-1.17. Centers for Disease Control and Prevention. 1992.

9. Frazier EL, Franks AL, Sanderson LM. Behavioral Risk Factor Data. In Using chronic disease data: A handbook for public health practitioners, pp 4.1-1.17. Centers for Disease Control and Prevention. 1992.

10. Groves RM. Survey Errors and Survey Costs. New York: John Wiley and Sons, 1989; 265, 271-272.

11. Levy PS, Lemeshow S. Sampling of Populations: Methods and Applications. New York: John Wiley and Sons, 1991; 347-350