

COMPARABILITY OF DATA (1998)

The BRFSS is a cross-sectional surveillance survey currently involving 52 reporting areas. 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 1998 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 question wording and a more restrictive age range for the HIV/AIDS questions. Idaho did not follow the protocol specifying that at least one call back attempt be made on any initial refusal or initial termination within interview. The following section identifies other known variations for the 1998 data year.

A. 1998 Data Anomalies and Deviations from Sampling Frame and Weighting Protocols

Alaska

Alaska's sample design is made up of three random digit-dialed strata and four listed strata, where phone numbers are selected from a list of putative household numbers. The sampling frame contains all of the household numbers in the random digit-dialed strata and an estimated 45% of all household numbers in the listed strata. This represents an exclusion of approximately 9% of all residential numbers in Alaska.

In 3 out of 4 geographic strata, prefixes are assigned to a high (RDD) or low (listed) density stratum depending largely on the number of active household numbers (about 2000 being the threshold) in the exchange (10,000 numbers). All numbers in the fourth geographic stratum are assigned to the listed density stratum. In the RDD strata, the probability that a number is selected depends on the number of active household numbers in its exchange. In the RDD strata, a number of prefixes equal to the target sample size are selected at random in proportion to their number of active HH's. (A prefix can be selected more than once.) Once a prefix is selected into the RDD sample, 48 suffixes are randomly generated. Once the entire sample for the stratum is generated duplicate numbers are deleted without replacement. The final result is h lists of phone numbers per stratum, each containing 48 or fewer unique numbers, where h is the target number of completes per stratum. Phone numbers are called sequentially from each list until one complete per list is obtained.

California

California, in July through December 1998, used a sample design in which only phone numbers from hundred blocks with one or more listed household numbers were included in the sampling frame. Such hundred blocks are estimated to contain 97.8% of all household numbers in California. Numbers from this frame were selected with an equal probability of selection.

Hawaii

Hawaii uses a sampling frame obtained from the state's telephone company. The frame purportedly contains all working prefixes in Hawaii. The frame is divided into six strata, corresponding to the major islands of Hawaii. The working prefixes are used to randomly generate telephone numbers with a probability proportional to the number of known household numbers in the hundred block to which a number belongs. Telephone numbers from hundred blocks with no known household numbers have a zero probability of being included in the sample. The percentage of telephone households excluded from the sampling frame is unknown. The numbers generated are subsequently screened by GTE Hawaii, who screens out all but the listed and unlisted household numbers in the sample.

Other areas

In a few states, a portion of sample records intended for use during one month may have been completed in another month. This deviation should only affect analyses based on monthly, rather than annual, data.

B. Other 1998 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 2.0% non-coverage in Utah, to 13.6% in Arkansas. It is estimated that 24% of households in Puerto Rico are without telephones.

Dual questionnaires and/or partial year coverage occurred in Illinois and Tennessee. Illinois used dual questionnaires and collected data on core items involving exercise, cigar smoking, fruits and vegetables and weight control and modules concerning immunization, cholesterol, hypertension, colorectal screening, injury and alcohol consumption for only six months of the interviewing period. Tennessee collected data for the modules on sexual behavior, family planning, injury and alcohol consumption for only part of the year. Data users will need to alter program code so that the usual "missing/dk/refused" codes are not combined with "9's" appearing in records due to noncoverage in the states mentioned here.

Almost all states define a completed interview as one in which information for the respondent is available through at least the demographic section of the questionnaire. However, Idaho, Montana, Oregon, Texas and Wyoming require completion of all questions on the core, modules, and state-added sections for an interview to be given the disposition of "complete". California modified the wording of mammography, Papanicolaou (PAP) smear questions and chronic alcohol use. These questions may have limited comparability to those of other reporting areas. California also asked the HIV/AIDS section questions to persons 18-45 years of age rather than to those 18-64 years of age as specified.

Almost all states collect data monthly, but a few states missed one or more months of collection in 1998. Maryland did not collect data in January, October, November and December. Hawaii and Maine did not collect data in January. Oregon did not collect data in January and February.

Income non-response varies substantially by reporting area. Although the median non-response rate on this item is 6.6%, it ranges from 0.9% to 25.4% across states. More than twenty-five percent of Arizona's respondent records were missing income, as were 20% of the records in Oklahoma. Compared with other items on the survey, income non-response is relatively high. For example, item non-response for age has a median of less than 0.3%, and a maximum of 1.5%. Other demographic items (not shown in the tables), including education, employment, and marital status have less than 1% item non-response.

Compare_98.rtf