

Sampling and Oversampling in the National Health Interview Survey

Chris Moriarity (cdm7@cdc.gov)

National Center for Health Statistics, Centers for Disease Control and Prevention

3311 Toledo Rd, Room 2209, Hyattsville, MD 20782

Abstract

The National Health Interview Survey (NHIS) is a continuous survey that collects health data using personal interviews. Changes have been made in the sample design for the NHIS (a "redesign") every 10 years or so, using information from the previous decennial census. The 2006 NHIS is the first year of the most recent sample redesign. Beginning with the previous NHIS design, implemented in 1995, and continuing with the 2006 design, minority persons are oversampled using two mechanisms: variation in sampling rates in substrata, and screening. Primary sampling units are partitioned into substrata based on population proportions of minority persons, and substrata with higher minority population proportions are sampled at a higher rate. A proportion of the sample cases are "screened"; that is, the household interview stops if no eligible minority persons are on the household roster.

Keywords: Sample Survey, Screening

1. Introduction

The National Health Interview Survey (NHIS) is the principal source of information on the health of the civilian noninstitutionalized population of the U.S. The NHIS sponsor is the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention. The NHIS is a continuous survey that conducts interviews at approximately 35,000 living quarters (households and noninstitutional group quarters such as college dormitories) each year. All eligible (i.e., civilian) persons at a sampled address are included in the NHIS interview, yielding a sample of approximately 87,500 persons each year. Each interview is conducted via a personal visit to the living quarters by an employee of the U.S. Bureau of the Census, which is the data collection agent for the NHIS.

The NHIS sample consists of clusters of living quarters chosen within a first-stage sample of U.S. counties. This sampling method is used to control the costs related to personal visit interviewing. The cost of conducting personal visit interviews in a simple random sample of U.S. living quarters would be prohibitive, due to the amount of travel that would be required.

A new sample redesign of the NHIS was implemented in 2006. The new sample redesign is very similar to the previous sample design, which was in effect from 1995 to 2005. The procedures used in the previous sample design to oversample black persons and Hispanic persons are still part of the current sample design. These procedures are described in more detail below. Changes that were made in the procedures for the new sample design also are described.

2. Oversampling Methods Used in the 1995-2005 and 2006- NHIS Sample Designs

Two main strategies are used to oversample selected subgroups of the U.S. civilian noninstitutional population for the NHIS:

1. "Substratification": Each primary sampling unit (one or more adjacent counties) that is selected into sample is partitioned into substrata, based on population proportions of the subgroups to be oversampled. Differential sampling rates are used in the substrata to increase the number of persons selected from the targeted subgroups.

2. "Screening": The initial sample is partitioned into two groups prior to interviewing. All households in the first group are interviewed with certainty. In the second group, the interviewer starts each interview with the collection of the household roster, as is done for a regular interview. For the 1995-2005 design, interviewing in the second group continued past the household roster enumeration only if one or more eligible black or Hispanic persons were on the roster; all other households were "screened out". For the new sample design that was implemented in 2006, interviewing in the second group continues past the household roster enumeration only if one or more eligible black, Asian, or Hispanic persons are on the roster; all other households are screened out.

More details are provided below about these strategies.

3. Substratification

Substratification is a commonly-used method in survey sampling for increasing the sample size of subgroups of

interest. The primary sampling units for the NHIS are counties (which includes county equivalents such as parishes in Louisiana, independent cities in Maryland, Missouri, Nevada, and Virginia, etc.) or aggregates of adjacent counties. The primary sampling units typically have large, heterogeneous populations, and so they are not assigned to sampling strata on the basis of population proportions of black, Hispanic, or Asian persons. Instead, the primary sampling units are split into substrata using information from the previous decennial census. For both the 1995-2005 design and the current design, 20 race/ethnicity density substrata were defined; however, there are some differences in the substrata definitions between the two designs, which are described below.

For both the 1995-2005 design and the current design, there is one additional "new construction" or "permit" (building permit) substratum. No oversampling occurred in this substratum in either the 1995-2005 design or in the current design, so this substratum is not discussed any further in this paper.

Extensive detail of the substrata for the 1995-2005 NHIS design is available elsewhere (Botman, et al. (2000), page 13, Table 4), so only summary-level information is provided here. A condensed version of Table 4 from Botman, et al. (2000) is Table 1. 1990 Decennial Census data were used to define substrata based on cross-classifications of the percentage of black persons (e.g., less than 10 percent, 10 percent up to 30 percent, etc.) and the percentage of Hispanic persons (e.g., less than 5 percent, 5 percent up to 10 percent, etc.). Over half of the 1990 U.S. population fell into the "lowest minority density" substratum defined as less than 10 percent black and less than 5 percent Hispanic. No other substratum had more than 8 percent of the 1990 U.S. population.

The substrata for the 2006- NHIS design are defined somewhat differently than the 1995-2005 substrata (Parsons, et al. (2003)). The most important difference is that the percentage of Asian persons is part of some of the substrata definitions. Also, population information from the 2000 Census was used, instead of 1990 Census data. Another new feature is that the substrata cross-classification boundaries were allowed to vary slightly from primary sampling unit to primary sampling unit. More details are forthcoming in an NCHS report that is scheduled to be completed in 2007. In summary, the substrata for the 2006- design can be described conceptually in terms of Census 2000 population distributions of Hispanic persons (H), Non-Hispanic black (B) persons, and Non-Hispanic Asian (A) persons as follows:

"low", "medium HBA", "medium H", "medium HB", "medium B", "medium A", "medium HA", "mixed HBA", "mixed H", "mixed HB", "mixed B", "mixed A", "mixed HA", "high H", "higher H", "highest H", "high B", "higher B", "highest B", and "high A".

Analogous to the 1995-2005 design, more than half of the 2000 U.S. population is in the "low" substratum, and none of the other substrata have more than 8 percent of the 2000 U.S. population.

Judkins, et al. (1999, page 100) discuss an issue associated with substrata definitions that are static, by necessity, for approximately a decade, because the sample for the entire design period is drawn at the beginning of the design. Population shifts can be expected to occur during the sample design period. For example, Judkins, et al. (1999) found a substantial migration of black persons had occurred between 1980 and 1988 out of areas with strong black concentrations in 1980. It is possible that substratum definitions that lead to efficient oversampling of targeted groups early in the design period could deteriorate and be much less efficient later in the design period due to population shifts. Limited options are available for making adjustments in the sampling allocations to the substrata, and there would be a lag time before the adjustments could be implemented.

NCHS has not yet done extensive research on the population shifts that occurred in the substrata during the 1995-2005 NHIS design, although this research is planned. Preliminary research by Taylor (2006) indicates that some noteworthy population shifts occurred in the 1995-2005 substrata, analogous to what Judkins, et al. (1999) observed: some significant movement of black and Hispanic persons by 2005 out of areas that had substantial black and/or Hispanic concentrations in 1990. The findings from this research, once completed, will come too late to influence the substrata definitions for the 2006- NHIS design, but they will be available for use when planning begins for the next NHIS sample design. They will be available in time to provide background information to help decide whether sampling allocation adjustments should be considered for sometime around the middle of the current design.

4. Screening

Screening is not as common in survey sampling as substratification. Screening was a major change in the NHIS sample design that was introduced in 1995, after extensive research (Judkins, et al. (1999)) suggested that it was a cost-effective way to meet the design goals of oversampling black and Hispanic persons while

controlling the overall sample size. Areas with high concentrations of black and/or Hispanic persons could be oversampled, and the use of screening would allow some of the households in those areas with no black or Hispanic persons to be eliminated from the interviewed sample, thus helping to control the overall sample size. The use of screening, coupled with careful attention paid to assigning variable oversampling rates and variable screening rates in the substrata, can achieve oversampling goals while controlling the amount of variability in the sampling weights, a desirable feature for a robust sample design.

For both the 1995-2005 design and the current design, the screening procedure is implemented in the same way. In a given substratum, prior to interviewing, the total sample is randomly partitioned into two pieces. The proportions assigned to the two pieces vary from substratum to substratum, depending on the nonminority household subsample rate (refer to Table 1). For example, substratum 1 for the 1995-2005 design used a 70.32% - 29.68% split, substratum 2 used a 46.88% - 53.12% split, etc. All households in the piece corresponding to the substratum nonminority subsample rate were interviewed with certainty. In the other piece, the screening procedure was applied, and only those households with one or more eligible black or Hispanic persons were retained for a full interview. For the current design, the same procedure is occurring, with the only change being that when the screening procedure is applied, only those households with one or more eligible black, Hispanic, or Asian persons are retained for a full interview.

The reciprocal of the substratum nonminority subsample rate is applied as an additional weight factor to nonminority households retained in sample in the substratum. If sufficient resources are available in the next few years, NCHS plans to investigate whether observed subsample rates varied substantially from expected subsample rates. If so, research should be carried out to determine if it is better to use the observed subsample rate to create the additional weight factor, rather than the expected subsample rate.

A key parameter when considering the use of screening in a survey design is the cost of an abbreviated interview that ends with the household being screened out (a "screen-out interview"), relative to the cost of a full interview. If the two costs are almost identical, screening provides little or no potential gain for a fixed overall survey budget. If a screen-out interview is thought to be much less expensive than a full interview, a substantial amount of screening can be justified to achieve large-scale oversampling of selected subpopulations. When planning began for the 1995-

2005 NHIS design, this parameter was an unknown quantity, due to lack of recent previous experience. The 1995-2005 NHIS design planners assumed that the cost of a screen-out interview was 1/3 of the cost of a full interview. The 2006- NHIS sample design researchers revisited this assumption. Although the survey designers reached a consensus that the cost of a screen-out interview was more than 1/3 of the cost of a full interview, due primarily to the fact that the same amount of overhead expense is required to get the interviewer to the living quarters to begin any type of interview, no consensus was reached on a revised estimate. The 2006-NHIS design assumed that the cost of a screen-out interview was 1/2 of the cost of a full interview. Although this assumed value is an upward adjustment of the previous assumed value, I conjecture it still is too low. Jones, et al. (2002) presented sample design research results where the assumed cost of a screen-out interview was 80% of the cost of a full interview.

If screening is going to be considered as part of the next NHIS sample redesign, it is important to obtain a good estimate of this key parameter early on, along with an assessment of the level of uncertainty in the estimate, to aid in the decision about the level of screening to be used (if at all).

5. Anticipated Sample Composition Changes for the 2006- NHIS

The average anticipated annual sample size for the 2006-NHIS is slightly smaller than for the previous design. The current sample design was done under the assumption of a fixed budget, which required a sample cut to absorb cost increases due to inflation. (We anticipate the annual number of completed interviews for the 2006- NHIS to be slightly smaller than the previous design, as well.) The sample design goals for the 2006- NHIS included maintaining approximately the same annual sample sizes of black persons and Hispanic persons as the previous design, and increasing the annual sample size of Asian persons. We anticipate the annual sample sizes of black persons and Hispanic persons to be similar to those of the 1995-2005 NHIS design.

We will not know for sure until the 2006 data collection is complete what the annual sample size will be for Asian persons, but we anticipate a substantial increase. Research conducted for the 2006- NHIS design suggested that just expanding the screening protocol to retain households with Asians would increase the annual Asian sample size by approximately 75% (Jones, et al. (2002)). The effect of defining some substrata using Asian population proportions will not be known until the 2006 NHIS data are available.

Preliminary results from the first quarter of 2006 NHIS data show an increase in the proportion of Asians in the NHIS sample relative to 2005, from less than 4% to more than 5%.

Maintaining the same annual sample sizes of black and Hispanic persons, while increasing the sample of Asian persons and reducing the overall sample size, implies that the sample size for persons who are not black, Asian, or Hispanic will decrease. However, persons who are not black, Asian, or Hispanic will still constitute a majority of the total sample, and we do not anticipate a substantial loss of precision for this group's sample-based estimates.

6. Conclusions, Summary

Research for the 2006- NHIS design showed that the two strategies employed for oversampling black and Hispanic persons in the 1995-2005 NHIS design were effective, and thus these strategies were retained for the 2006- NHIS design.

Substrata definitions are fixed for a sample design period that lasts approximately ten years. Population shifts during the sample design period can cause inefficiencies to occur in the later years of the period. Research is planned by NCHS to study the effect of population shifts during the 1995-2005 NHIS sample design. The results from this research may influence how substrata are formed for the next design. The research findings may also suggest that planning should be done to allow for sample allocation adjustments somewhere around the midpoint of the current design.

If screening is going to be considered as part of the next NHIS sample redesign, it is important to obtain a good estimate early on of the cost of a screen-out interview relative to a full interview, along with an assessment of the level of uncertainty in the estimate. This information will allow for an informed decision about whether screening is to be used, and if so, at what level.

References

Botman S, Moore T, Moriarity C, and Parsons V. Design and Estimation for the National Health Interview Survey, 1995-2004. *Vital Health Stat 2(130)*. 2000

Jones C., Moriarity C., Davis K., and Gonzalez J. Methods to Improve the Precision of Health Statistics for Non-Hispanic Asians in the National Health Interview Survey, *ASA Proceedings of the Joint Statistical Meetings, American Statistical Association, 2002, 1695-1700.*

Judkins D, Marker D, Waksberg J, et al. National Health Interview Survey: Research for the 1995-2004 Redesign. *Vital Health Stat 2(126)*. 1999.

Parsons V., Katzoff M., and Zha W. Within-PSU Sampling Strategies for the Redesigned National Health Interview Survey, *ASA Proceedings of the Joint Statistical Meetings, American Statistical Association, 2003, 3203-3207.*

Taylor, L. Year-to-year Change in the NHIS Sample Composition for the Years 1995-2005. Unpublished internal NCHS report, 2006.

Table 1: 1995-2005 NHIS race/ethnicity density substrata (excerpt of Table 4 in Botman, et al. (2000), page 13)

Substratum	1990 Census density: % black	1990 Census density: % Hispanic	Percent of U.S. population in 1990	SSU oversampling rate	Nonminority household subsample rate
1	<10	<5	55.3	1	0.7032
2	<10	5-10	7.7	1.5	0.4688
3	<10	10-30	8.0	1.5	0.4688
4	<10	30-60	3.1	1.6	0.4395
5	<10	60+	3.5	2.3	0.3057
6	10-30	<5	4.6	1	0.7032
7	10-30	5-10	1.4	1.5	0.4688
8	10-30	10-30	1.9	1.5	0.4688
9	10-30	30-60	0.9	1.5	0.4688
10	10-30	60+	0.7	2	0.3516
11	30-60	<5	2.3	1	0.7032
12	30-60	5-10	0.5	1	0.7032
13	30-60	10-30	0.8	1	0.7032
14	30-60	30-60	0.7	1.5	0.4688
15	30-60	60+	0.2	2	0.3516
16	60+	<5	6.6	1.05	0.7032
17	60+	5-10	0.6	1	0.7032
18	60+	10-30	0.9	1.2	0.5860
19	60+	30-60	0.2	1.5	0.4688
20	60+	60+	<0.05	1.5	0.4688
Permit	N/A	N/A	N/A	1	1

Notes:

For intervals of form "x-y" in columns 2 and 3, the lower endpoint is included and the upper endpoint is not included.

SSU: "secondary sampling unit"

"Nonminority" includes everyone except black and Hispanic persons.

"Permit" denotes the "building permit" or "new construction" substratum. No oversampling or screening occurred in this substratum.

An SSU oversampling rate of "1" denotes no oversampling. An SSU oversampling rate greater than 1 indicates oversampling. See Botman, et al. (2000) for details of the SSU oversampling procedure.

A nonminority subsample rate of "1" denotes no subsampling of nonminority households. A minority subsample rate less than 1 is the expected retention rate of nonminority households after the screening procedure is applied.