

METHODS FOR CALCULATING NONRESPONSE FOR "SCREENER HOUSEHOLDS" IN THE NHIS¹

Frances M. Chevarley, National Center for Health Statistics
Frances M. Chevarley, NCHS, Room 850, 6525 Belcrest Road, Hyattsville, MD 20782

Key Words: nonresponse rate, health survey, sample design

1. Introduction

In order to produce more reliable estimates for Black and Hispanic persons in the National Health Interview Survey (NHIS), more households with Black and Hispanic persons are selected by its 1995-2004 sample design than other types of households. This is done by oversampling in areas with higher proportions of Black and/or Hispanic persons, and also by designating some households as "screener households". The "screener households" are interviewed only when they contain Black and/or Hispanic residents and are not interviewed otherwise. This is the first time that "screener households" have been used in the NHIS.

Because response rates are critical in survey management, and also play a critical role in making nonresponse post-survey adjustments to compensate the bias of survey estimates, it is important to explore how "screener households" may affect the calculation of nonresponse. The purpose of this paper is to describe several definitions used for calculating nonresponse rates when there are "screener households" and to suggest further research.

2. Background

The National Health Interview Survey (NHIS) is a nationally representative survey of the civilian, noninstitutionalized population of the United States. Since 1957, the National Health Interview Survey has been conducted annually and has been sponsored by the National Center for Health Statistics (NCHS). Personal interviews are conducted in sample households by specially trained interviewers employed by the Census Bureau.

NHIS includes a set of "core" questions that are changed every 10 years or so. The basic questionnaire collects information on health status and health care utilization. In addition, the NHIS includes special supplements on special topics for a given year.

3.1 1995-2004 NHIS Sample Design

The NHIS sample design is redesigned or modified every 10 years after the national Census is taken. As reported by NCHS, 1998, and also by Westat, 1996, under a contract with NCHS, the latest sample design (1995-2004) was based on the 1990 census and consists of about 42,000 interviewed households containing about 106,000 persons.

A Primary Sampling Unit (PSU) in the NHIS consists of a county, a small group of contiguous counties, or a Metropolitan Area (MA). About 1,900 PSU's were defined from the more than 3,000 counties or equivalents covering the U.S.

First the PSU's for the 52 largest metropolitan areas were defined to be self-representing strata. Within each State the remaining PSU's were combined into strata. As part of this process an additional 43 PSU's were also designated as self-representing strata. Thus there are a total of 95 self-representing strata. The NHIS PSU's are combined into a total of 237 strata, 95 of which are self-representing.

In each non self-representing stratum usually two PSU's are selected for the NHIS sample; in 21 strata with small population sizes one PSU was selected for the NHIS sample. This resulted in a total of 358 PSU's selected for the sample.

Within a PSU, two types of second stage units (called segments) are used: area segments and permit area segments. Area segments are defined geographically and contain an expected eight or twelve households. Permit area segments cover geographical areas containing housing units built after the 1990 census.

In each sample PSU, up to 20 substrata for the area sample were formed based on a cross-classification of the proportions of the population that was Black and Hispanic. A separate substratum is used for the permit frame of housing units built since the 1990 Census. Segments (or second-stage sampling units) in substrata with higher proportions of Black and Hispanic populations were selected with higher probabilities for the NHIS sample.

In addition to oversampling segments, NCHS also screens for Black and Hispanic persons in all of the area segments. This procedure has been carried out by

¹The views expressed in this paper are those of the author and no official endorsement by the Department of Health and Human Services, or the National Center for Health Statistics is intended or should be inferred.

designating a subsample of the addresses in each segment to be retained in the survey only if they contain a Black and/or Hispanic person. This allows for better control of the variation of the number of sample households in the segments.

Each year about 72,000 addresses are designated for the usual NHIS sample; when we take into account the 20 percent of addresses expected to be demolished or unoccupied units, the NHIS annual sample yields about 57,000 households, some of which are designated as "screeener" households. Taking into account that some of the households will be screened out, and the NHIS expected 5 percent non-response rate, the NHIS sample is expected to include data on about 42,000 households, containing about 106,000 persons.

Table 1

Year	1995	1996	1997
Sampled Households	39,239	24,371	40,297

As shown in Table 1, the 1995 NHIS sample consisted of 39,239 interviewed households containing about 102,000 persons. This is less than the 42,000 expected number because the 1995 NHIS was only in the field for 48 weeks and not the usual 51 weeks. One week was lost because it was used for the annual interviewer training and an additional three weeks were lost in 1995 due to a government shutdown. The reason the sample is small in 1996 is that an abbreviated version of the new CAPI questionnaire was tested with part of the NHIS sample in 1996. The new NHIS CAPI questionnaire went into the field in 1997. The number of interviewed households is higher in 1997 than in 1995 and 1996. However it is still slightly below the expected 42,000 number because two weeks instead of the usual one week were used in interviewer training and the non-response rates in 1997 were higher than expected.

3.2 1995-2004 and 1985-94 sample designs compared

The 1995-2004 NHIS sample design compared with the 1985-94 sample design has more PSU's (Westat, 1996), (NCHS, 1989). Its stratification was by 50 States and DC compared with stratification by 4 regions in the previous sample design. There are more designated housing units for the newer design, however some of these are designated as screening housing units. The expected interviewed households for the 1995-2004 design (42,000) is less than before (49,000) and the expected number of interviewed persons (106,000) is also less than before (120,000).

4.1 NHIS Valid Outcome Codes

Type A households are eligible households in which the interview was not completed. There are five valid codes for Type A noninterviews for the NHIS: refusals; no one home, repeated calls; temporarily absent; language problem; and other. Unlike Type A's, Type B and Type C households are not eligible and deal with such situations as vacant units, occupied entirely by persons whose usual residence is elsewhere, occupied entirely by armed forces members, demolished, outside segment boundaries, etc.

Starting with the 1995-2004 sample, a Type B code was added for screening cases that had been screened out because the household did not contain a Black and/or Hispanic member.

4.2 Final Outcomes for the 1995 NHIS

Table 2 shows the final outcomes for the 1995 NHIS sample. In 1995, the NHIS had a total of 36,309 households that were designated as "interview" households and that were eligible for interview (does not include the Type B's or C's). Out of these 34,064 were interviewed and 2,245 were Type A's.

There were a total of 19,444 households that were designated as "screeener" households; that is, the household was interviewed only if it contained a Black and/or Hispanic member. Of these there were 5,175 completed interviews; 13,929 that were screened out because these households did not contain any Black and/or Hispanic members; and 340 were Type A's.

Table 2

Final Outcomes for the 1995 NHIS			
Outcome	Interview HH	Screeener HH	TOTALS
Interviewed	34,064	5,175	39,239
Screened Out	0	13,929	13,929
Type A's	2,245	340	2,585
TOTALS	36,309	19,444	55,753

5. Methods

The usual way of calculating non-response is to divide the number of Type A's by the number of households eligible to be interviewed, that is by dividing by the Type A's plus those cases that were interviewed.

$$\text{non-response} = \frac{A}{I+A} \quad (1)$$

This definition is consistent with the first two response rates defined in the publication *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for RDD Telephone Surveys and In-Person Household Surveys* by The American Association for Public Opinion Research (AAPOR, 1998).

Below are two nonresponse definitions each of which provide specifications for the interview households, the screener households, and the combination of the interview and screener households. These two definitions are examples among many that could be defined when there are "screener households". Each of the two nonresponse definitions below use equation (1) in defining nonresponse for the interview households; other equations are specified for the screener households. All two definitions combine the information from the interviewer and the screener households in producing the nonresponse for the combination.

5.1 Calculating NRR1

For this first nonresponse definition, Nonresponse 1 (NRR1), equation (1) is used for the interview households, and for the screener households. For the combination of the interview and screener households, interview and Type A information is combined from the interview households and the screener households.

The calculations for NRR1 are shown below:

NRR1 for Interview Households:

$$NRR1_I = \frac{2,245}{34,064 + 2,245} = 6.2\%$$

NRR1 for Screener Households:

$$NRR1_s = \frac{340}{5,175 + 340} = 6.2\%$$

NRR1 for Combination:

$$NRR1_c = \frac{2,245 + 340}{34,064 + 5,175 + 2,245 + 340} = 6.2\%$$

For the combination, the 34,064 + 5,175 is the sum of the households that were interviewed for the interview and

screening households, respectively, and the 2,245 + 340 is the sum of the Type A's for the interview and screening households respectively. Although NRR1 rounded to 6.2% for the interview households, screener households, and combination, they were not exactly the same; it is unusual and unexpected for them to round to the same number.

5.2 Calculating NRR2

Equation (1) is used for calculating Nonresponse Rate 2 (NRR2) for the interview households. The following equation is used for the screener households:

NRR2 for Screener Households:

$$NRR2_s = \frac{eA}{I+eA}$$

where e is the estimated proportion of the Type A households that would contain Black and/or Hispanic persons. This equation is consistent with response rates 3 and 4 defined by AAPOR in their 1998 publication (AAPOR, 1998).

Using equation (1) for the interview households and the above equation for the screener households we get the following calculations for NRR2:

NRR2 for Interview Households:

$$NRR2_I = \frac{2,245}{34,064 + 2,245} = 6.2\%$$

NRR2 for Screener Households:

$$e = \frac{5,175}{5,175 + 13,929} = .271$$

where information in Table 2 is used in estimating the proportion 'e' of Type A households eligible to be interviewed. For the screener households we assume that the households that are interviewed were correctly interviewed and contain a Black and/or Hispanic person, and those that were screened out were correctly screened out and did not contain a Black and/or Hispanic person. We estimated 'e' by using the proportion of those interviewed from the interviewed and screened-out households.

In this case:

$$eA = .271 * 340 = 92.372$$

and

$$NRR2_s = \frac{92.372}{5,175 + 92.372} = 1.8\%$$

NRR2 for Combination:

$$NRR2_c = \frac{2,245 + 92.37}{34,064 + 5,175 + 2,245 + 92.37} = 5.5\%$$

6. Trends in nonresponse using these formulas

Nonresponse rates for the interview households, screener households and the combination were calculated for 1995-97 using these two definitions and are shown in Figure 1. Because the two definitions (NRR1, and NRR2) produce identical nonresponse rates for the interview households, there is only one trend line for the interview households in Figure 1. For the screener households, the trends look very different depending on which of the two definitions is used. The combination nonresponse rates, which combines information for the interview and screener households, is therefore a "total" nonresponse rate and is graphed with the total nonresponse rates for the earlier period 1986-94 before the NHIS had "screener households". These trends for the combination do show some differences depending on which of the two definitions is used.

7. Conclusion

In conclusion, the 1995-2004 NHIS sample design is the first NHIS sample design in which households are screened, specifically for Black and Hispanic persons. In this sample design some households were designated as "interview households", and were always interviewed. Others were designated as "screener households" and were interviewed only if they contained Black and/or Hispanic persons. When the sample design has "screener households", the calculation of nonresponse is more complicated. In this paper we defined two nonresponse rates. The only difference in the two nonresponse rate definitions was their respective assumptions about the Type A households for the screener households. For NRR1, all of the Type A's were used in calculating nonresponse for the screener households. For NRR2, an estimate of the proportion of the Type A's that would contain a Black and/or Hispanic member was made. The estimated proportion was based on the information about the number of screener households that were interviewed versus screened out.

The nonresponse rates for these two definitions (NRR1, and NRR2) are exactly the same for the interview

households (6.2%); and they vary greatly for the screening households (6.2%, and 1.7%, respectively). These values vary moderately for the combination of the interview and screening households (6.2%, and 5.6%, respectively). The values are largest for NRR1 and smallest for NRR2. However, other additional nonresponse estimates can also be defined depending upon other assumptions and information about the number of Type A's for the screener households.

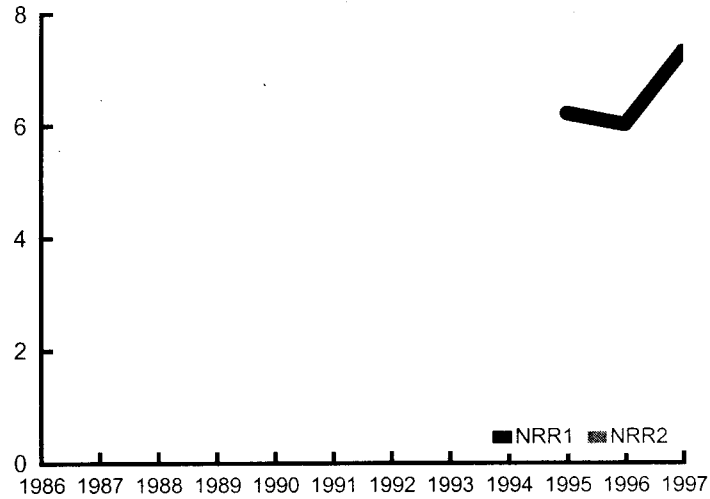
Our further study would include defining other additional nonresponse estimators and looking at their trends. We may also want to consider whether additional information about the Type A's could either be collected or could be extracted from existing survey data that would be useful in defining other nonresponse estimators.

Nonresponse rates don't only play a critical role in monitoring interviewer and survey performance, but they are also critical in making post-survey adjustments to compensate the bias of survey estimates resulting from nonresponse. As part of evaluating the different nonresponse estimators, our future research are not only to consider the implication of the different nonresponse estimators for monitoring interviewer and survey performance but also for producing post-survey adjustments, survey estimates based on the post-survey adjustments, and variances of the survey estimates.

REFERENCES

- The American Association for Public Opinion Research. 1998. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for RDD Telephone surveys and In-Person Household Surveys*. Ann Arbor, Michigan: AAPOR.
- National Center for Health Statistics. (1998). Benson V, Marano MA. Current estimates from the National Health Interview Survey, 1995. *Vital Health Stat* 10(199). 1998. DHHS Pub. No. (PHS)98-1527. Public Health Service. Washington: U.S. Government Printing Office.
- National Center for Health Statistics. (1989). Massey JT, Moore TF, Parson VL, Tadros W. Design and estimation of the National Health Interview Survey, 1985-94." *Vital and Health Stat*. 2(110). DHHS Pub. No. (PHS) 91-1386. Public Health Service. Washington: U.S. Government Printing Office.
- Westat. (1996). Judkins D, Marker D, Waksburg J, and Other Contributors. National Health Interview Survey: Research for the 1995 redesign. Report under NCHS Contract no. 200-89-7021. March 8, 1996.

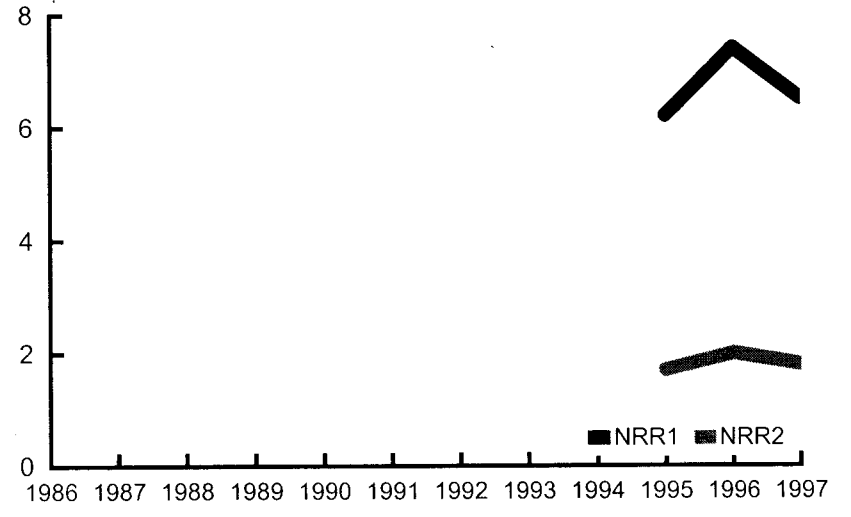
Interview Households



SOURCE: CDC/NCHS

Figure 1

Screener Households



SOURCE: CDC/NCHS

Combination



SOURCE: CDC/NCHS