

A DETAILED LOOK AT COVERAGE IN THE THIRD NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY (NHANES III, 1988-1994)

Jill M. Montaquila, Leyla Mohadjer, Joseph Waksberg, Meena Khare
Jill M. Montaquila, Westat, Inc., 1650 Research Blvd., Rockville, MD 20850

KEY WORDS: Undercoverage, Coverage, Survey Quality

1. Introduction

An important objective in the design of any sample survey is to adequately represent, or cover, the population of interest. Undercoverage generally results in biased estimates, and inferences may be flawed. While overcoverage is rarely a concern for survey practitioners, undercoverage is frequently an important consideration, particularly because certain demographic groups (e.g., black adult males) tend to be underrepresented in survey efforts. This paper examines coverage in The Third National Health and Nutrition Examination Survey (NHANES III). The sample design for NHANES III is summarized in Section 2. Section 3 describes the estimation of survey coverage. Comparisons between coverage of NHANES III and other national surveys are given in Section 4. In Section 5, comparisons of the coverage between the two phases of NHANES III are made. Conclusions are given in Section 6.

2. NHANES III sample design

NHANES III was a survey of health and nutrition characteristics of the total civilian noninstitutionalized population, 2 months of age or older, in the 50 states of the United States. A four-stage sample design was used: (1) Primary Sampling Units (PSUs) comprising mostly single counties, (2) area segments within PSUs, (3) households within area segments, and (4) persons within households.

The data collection for NHANES III was carried out between October 1988 and October 1994. In order to permit separate analyses for two 3-year periods (referred to as Phase 1 and Phase 2), as well as for the entire field period, the sample of PSUs was randomly allocated to two sets. One set was assigned to the first 3-year time period during which NHANES III was conducted (Phase 1, 1988-91), and the other set to the second 3-year period (Phase 2, 1991-94). The allocation of PSUs to the two phases was made in a way that retained as much of the original stratification as possible in each phase.

For most of the sample in Phase 1, the second stage of sampling was area segments comprising city or suburban blocks, combinations of blocks, or other area segments in places where block statistics were not produced in the 1980 census. The area segments were used only for a sample of persons who lived in housing units built before 1980. For units built in 1980 and later, the second stage of sampling consisted of sets of addresses selected from building permits issued in 1980 or later. (These are referred to as new construction segments.) In Phase 2, the 1990 census data were used for the selection of the second stage units, with no new construction sampling. For more detail on the NHANES III sample design, refer to NCHS (1992).

3. Estimating survey coverage

Poststratification was used for the same purposes in NHANES III as in many other household surveys, partly to reduce sampling errors and partly to dampen the effects of undercoverage. The latter goal is not fully accomplished unless the missed portion of the population in the survey has characteristics similar to the interviewed part, within the poststratification cells. This normally does not occur, so that poststratification generally reduces, but does not completely eliminate, coverage biases.

Survey coverage is generally estimated using the reciprocals of the poststratification factors. Separate coverage factors were computed for the two phases of NHANES III. The poststratification factors reflect sampling errors as well as coverage, which could distort measures of coverage for small sample sizes. Under such circumstances, it is usually reasonable to look for patterns rather than concentrating on coverage rates for individual groups.

Table 1 contains coverage rates for Phase 1, for Phase 2, and for the full 6-year sample. As noted in Section 2, the Phase 1 sample included new construction segments that consisted of dwelling units built after 1980. An additional overall poststratification step was applied to Phase 1 data to bring up the proportion of the sample residing in the newly constructed units to an associated census estimate; this adjustment is not reflected in the coverage rates. The coverage rates are thus appropriate

Table 1. NHANES III coverage rates¹

Age-race/ethnicity domain	Phase 1 coverage		Phase 2 coverage		Overall (Phases 1 and 2 combined) coverage	
Overall	91.5		91.5		91.5	
Under 19 years	94.0		96.7		95.4	
20 to 59 years	89.9		91.0		90.5	
60+ years	92.4		83.3		87.8	
	Male	Female	Male	Female	Male	Female
Overall	88.2	94.6	87.2	95.5	87.7	95.1
White/Other	87.6	93.9	85.1	94.2	86.3	94.1
2 to 11 months	81.8	84.9	89.3	93.3	85.4	89.1
1 to 2 years	82.4	96.3	83.1	102.3	82.8	99.3
3 to 5 years	83.6	84.7	87.7	101.8	85.7	93.3
6 to 11 years	86.3	89.4	88.3	94.0	87.3	91.7
12 to 19 years	98.1	106.6	83.4	109.0	90.7	107.8
20 to 29 years	89.0	90.0	88.5	97.0	88.7	93.4
30 to 39 years	83.7	98.9	74.7	95.4	79.1	97.1
40 to 49 years	83.9	94.7	94.4	91.5	89.4	93.0
50 to 59 years	83.4	85.9	91.8	89.5	87.7	87.7
60 to 69 years	92.3	82.8	81.0	89.0	86.7	85.9
70 to 79 years	90.9	105.5	72.6	78.6	81.4	91.8
80 years or more	93.4	97.4	91.0	88.7	92.1	92.8
Black, non-Hispanic	88.5	94.6	102.0	105.9	95.4	100.4
2 to 35 months	99.8	97.0	107.6	112.8	103.9	105.2
3 to 5 years	93.2	95.9	105.7	110.8	99.6	103.6
6 to 11 years	100.9	90.6	113.4	117.0	107.3	104.0
12 to 19 years	84.1	87.0	112.2	106.5	98.3	96.8
20 to 39 years	81.6	92.5	97.8	150.4	89.8	99.0
40 to 59 years	93.8	105.5	96.1	103.5	95.0	104.4
60 years or more	83.0	94.4	89.7	95.5	86.4	95.0
Mexican-American	95.7	105.3	86.1	91.2	90.7	97.8
2 to 35 months	120.1	106.8	97.3	107.9	107.4	107.4
3 to 5 years	115.5	112.2	98.9	88.7	106.6	99.8
6 to 11 years	98.3	116.3	97.4	101.3	97.8	108.4
12 to 19 years	101.5	106.0	94.8	93.8	98.2	99.7
20 to 39 years	86.4	102.4	82.0	96.8	84.1	99.5
40 to 59 years	91.9	105.4	73.2	70.7	81.6	86.4
60 years or more	95.0	87.3	70.5	69.4	82.0	77.7

¹Coverage rates reported are with respect to undercount adjusted independent estimates of totals obtained from the March 1990 (for Phase 1) and March 1993 (for Phase 2) Current Population Survey.

measures of the part of the U.S. population missed in the survey, and they can be considered guides to the total undercoverage. The coverage rates essentially reflect inadequacies in the implementation of the survey operations, *e.g.*, those caused by possible omissions of some housing units from the address listings, failure of respondents to report all household members, omission of persons with no fixed usual place of residence, etc. (A very small part of the undercoverage comes from the trimming operation in which the weights for a few sample persons were reduced.) These are the usual causes of undercoverage in household surveys.

4. Comparison of NHANES III coverage to that of other national surveys

The coverage ratios shown in Table 1 reflect undercoverage for all reasons, omitted addresses (including new construction units for Phase 1), missed persons, and trimmed weights. The overall coverage rates for the two phases are virtually identical. This is also the case for total males and total females. However, there are curious differences among the subdomains shown in Table 1. For whites, the coverage rates are reasonably consistent except for persons 60 years and over; Phase 1 coverage was higher than Phase 2 coverage for four of the six age-sex groups into which persons 60 years and over were divided. Black coverage was higher in Phase 2 for almost all age-sex groups, but the reverse held true for Mexican-Americans.

There was a small, but very consistent, downward trend in coverage with advancing age, from 95 percent for persons aged 0 to 19 years, to 91 percent for those aged 20 to 59 years, to 88 percent for persons 60 years and over. Coverage was particularly low for Mexican-Americans 60 years and over at 80 percent.

It is instructive to compare the NHANES III coverage rates with coverage in Census surveys. Tables 3 and 4 contain data for the CPS and the National Health Interview Survey (NHIS), two major household surveys conducted by the Census Bureau. It should be noted that NHANES III was poststratified to undercount adjusted totals. However, poststratification to undercount adjusted figures has been done only recently by the Census Bureau (1994 for CPS, 1995 for NHIS). Direct comparison to Census surveys in earlier time periods will therefore be misleading. Similarly, NHANES III should not be compared to earlier cycles of NHANES unless some adjustments are made. Table 2 contains estimates of undercoverage in recent

censuses which essentially account for the differences between the census-based counts and the undercount-adjusted figures.

NHANES III coverage compares favorably with that of major surveys. Tables 3 and 4 show coverage in CPS and NHIS for recent time periods. Overall coverage in CPS was 92.1 percent as compared to 91.5 percent in NHANES III. The NHIS coverage rates (for the third quarter of 1992, the latest period available) shown in Table 4 cannot be compared directly to the NHANES figures because at that time the NHIS poststratified the sample estimates to unadjusted census population counts. Table 2 indicates that an adjustment would reduce NHIS coverage by 1.8 percent. With that adjustment, NHIS coverage for the time period would be 89.1 percent, about 2 percent below NHANES III.

It is also illuminating to compare minority coverage rates in NHANES with those in Census surveys. The Census surveys have had a long history of serious undercounts of blacks and Hispanics, particularly males aged 20 to 49 years, and the coverage of these groups is frequently considered a more sensitive indicator of potential bias than the overall coverage rate. In the October 1994 CPS, coverage of blacks was about 83 percent with black males aged 20 to 49 years in the range 66.0 to 81.6 percent. NHIS coverage of blacks was even worse. If one adjusts the coverage rates in Table 4 to take into account the 1990 Census undercounts shown in Table 2, the black coverage rate in NHIS was 77 percent and the rate for black males aged 20 to 49 years was in the range 58 to 79 percent. NHANES III coverage for all blacks was 98 percent, and for black adult males aged 20 to 49 years coverage was about 93 percent. The NHANES III field experience in black households was quite good, and satisfactory levels of coverage were expected, but the coverage rates attained were surprisingly high, even higher than those seen in recent decennial censuses.

Coverage rates for blacks and for Mexican-Americans were higher than those for whites and persons of other race/ethnicity. This is virtually unheard of in household surveys. The NHANES III interviewers were thoroughly trained, experienced, and very closely supervised, and it is believed that was an important reason for the high coverage. Respondents, particularly those with low incomes who are disproportionately minority, were highly motivated to cooperate, partly due to the monetary incentive and partly due to the opportunity to get free medical

examinations. However, the effect on coverage of these incentives is likely to have been small. For the majority of cases, incentives were introduced after the household roster was obtained. In the few cases where additional persons were mentioned during the course of the interview, the added persons were not included in the sample, and did not affect the weights of the sampled persons.

The Mexican-American coverage was also quite high - 94 percent overall, 91 percent for males, and 98 percent for females. For young children some of the rates exceeded 100 percent, indicating overcoverage. Some of the reasons for the high coverage are probably the same as for blacks. However, Mexican-American coverage was also affected by another feature of the NHANES III operating procedures.

NHANES III (and Hispanic HANES) used a more inclusive definition of Mexican-Americans than is used by the Census Bureau. In carrying out the census and current surveys, the Census Bureau does not attempt to influence respondents' perception of their race, ethnicity, or country of origin. A considerable proportion of Hispanics in several Southwestern states do not ordinarily describe themselves as Mexican-Americans, particularly in New Mexico, Colorado, and Texas. These states have not had a large influx of Hispanics from other parts of Latin America. It is likely that many of the descendants of persons who lived in this area at the time the U.S. annexed the territory from Mexico do not consider their country of origin to be Mexico. However, for both HHANES and NHANES III these persons were classified as Mexican-Americans, and this issue was stressed in the interviewer training. Although there is no way of getting an exact estimate of the effect of the difference in classification, Census data indicate this redefinition probably increased the Mexican-American population by 3 to 4 percent.

The overcoverage of Mexican-American children implies that there is a possibility that non-poststratified Mexican-American estimates would provide more accurate statistics than the data that are poststratified to adjusted census data. However, some of the reasons for considering the NHANES III data to be more accurate than the census data were speculative. More importantly, a major reason for poststratification was to create consistency with other government data, and not doing so would cause complications for data analysts who might also be looking at vital statistics rates and numbers, data from the NHIS, etc. It is also useful to note that the analyses that use percentage

distributions or regression parameters rather than aggregate totals will be only trivially affected by poststratification.

Coverage for non-Mexican-American whites together with those of other races (excluding blacks) was about 90 percent, which is lower than coverage for blacks and for Mexican-Americans. The rates cannot be compared exactly to CPS or NHIS. CPS coverage is available only for non-Hispanic whites, as opposed to non Mexican-Americans, and NHIS coverage rates are not available for either all Hispanics or Mexican-Americans. Nevertheless, Tables 1, 3 and 4 indicate that NHANES III coverage of the white population was probably about the same as NHIS, and a little lower than CPS.

5. Comparison of coverage between the two phases of NHANES III

The Phase 1 and Phase 2 coverage rates shown in Table 1 for the total population are virtually identical. The rates for white persons are mostly similar for the two phases. Coverage is higher in Phase 2 for blacks and lower for Mexican-Americans, but coverage of minorities in both phases was still considerably better than household surveys typically achieve.

The similarity of the two sets of coverage rates is perplexing. Phase 2 did not include new construction segments and thus there was no need for a "new construction" poststratification adjustment in Phase 2. The rates for Phase 1, given in Table 1, reflect coverage prior to the new construction adjustment. Consequently, one would expect Phase 1 coverage to be considerably below Phase 2 coverage because presumably, Phase 1 suffered from the same problems as Phase 2, in addition to the coverage problems in the new construction segments. The data in Table 1 seem to imply there was a sudden and sizable reduction in coverage that occurred at the start of Phase 2. However, this is doubtful, since the same field staff implemented Phase 1 and 2, the training and supervision were unchanged, and the Phase 2 area maps were probably better than the ones available in Phase 1, due to the availability of the Census system of automated maps (TIGER maps).

A possible problem could be in the control totals used for poststratification. The 1990 CPS population totals were based on Census Bureau projections of the 1980 population, adjusted for census undercounts; the 1993 CPS data used projections from 1990. For at least some of the domains, the 1990 census counts were

not in line with the Census Bureau's demographic projections from 1980. This was particularly true for Hispanics, for whom the 1990 census counts exceeded the projections by a significant amount.

6. General observations

The Census Bureau's difficulties in projecting the population counts needed for control totals through the 1980's make the Phase 1 poststratification factors an uncertain measure of Phase 1 coverage. Phase 2 coverage is probably a more reliable indication of true NHANES III coverage. Phase 2 coverage can be summarized as follows:

1. Overall coverage in NHANES III was about a half percent below CPS but about a half percent over that in NHIS. The white, non-Hispanic coverage in NHANES III was also midway between CPS and NHIS.
2. Black coverage in Phase 2 of NHANES III was greater than 100 percent. It is not clear whether this was due to sampling error or somewhat different race coding. Possibly NHANES III interviewers followed a somewhat different practice in multi-racial households than the Census Bureau practice. The overcoverage applied to almost all sex-age groups except for adult males and both males and females 60 years and over. The coverage rate for black males 20 to 59 years of age, frequently considered the most sensitive indicator of the adequacy of coverage, was 97 percent, which is far above the 76 percent in CPS or the under 60 percent in NHIS.
3. Mexican-American coverage was 88 percent. It cannot be directly compared to CPS since the CPS poststratification was carried out for

total Hispanics and restricted to persons 16 years and over. However, since Mexican-Americans account for about two-thirds of all U.S. Hispanics, it is likely that Mexican-American coverage in CPS is fairly similar to that of all Hispanics. The CPS coverage for all Hispanics 16 years and over was 82 percent. The NHANES III coverage for the same ages was about 84 percent. The higher coverage in NHANES III held for both males and females.

4. There was a steady decline in coverage with advancing age, from 97 percent for persons under 19 to 83 percent for those 60 years and over. This is the reverse of the CPS pattern. The decline appears to be somewhat more pronounced for blacks and Hispanics, but it also exists for the white and other category.

REFERENCES

- National Center for Health Statistics (1992). "Sample Design: Third National Health and Nutrition Examination Survey." *Vital and Health Statistics*, Series 2, Number 113, DHHS Publication No. (PHS) 92-1387. Hyattsville, Maryland.
- Robinson, J.G., B. Ahmed, P. Das Gupta, and K. Woodrow (1993). Estimation of population coverage in the 1990 U.S. census based on demographic analysis. *Journal of the American Statistical Association* 88(423):1061-1079.
- U.S. Bureau of the Census (1995). Current Population Survey, October 1994: School Enrollment Technical Documentation. Washington: U.S. Bureau of the Census.

Table 2. Historical decennial census coverage rates by race and sex: 1940-1990

Race and Sex	1990	1980	1970	1960	1950	1940
TOTAL	98.2	98.8	97.3	96.9	95.9	94.6
Male	97.2	97.8	96.6	96.5	95.6	94.2
Female	99.1	99.7	98.0	97.3	96.2	95.0
Black	94.3	95.5	93.5	93.4	92.5	91.6
Male	91.5	92.5	90.9	91.2	90.3	89.1
Female	97.0	98.3	96.0	95.6	94.6	94.0
Non-Black	98.7	99.2	97.8	97.3	96.2	95.0
Male	98.0	98.5	97.3	97.1	96.2	94.8
Female	99.4	99.9	98.3	97.6	96.3	95.1

SOURCE: Robinson *et al.* (1993)

Table 3. CPS coverage rates¹, October 1994

Age	Non-Black		Black		All persons		Total
	Male	Female	Male	Female	Male	Female	
0 to 14 years	92.9	96.4	85.0	83.8	91.6	94.3	92.9
15 years	93.3	89.5	76.3	82.4	90.5	88.3	89.5
16 to 19 years	88.1	89.1	71.1	80.2	85.5	87.7	86.6
20 to 29 years	84.7	89.7	66.0	81.1	82.3	88.4	85.4
30 to 39 years	90.4	93.1	68.0	84.5	87.7	92.0	89.9
40 to 49 years	92.8	96.6	81.6	91.1	91.7	95.9	93.8
50 to 59 years	95.3	97.4	89.6	92.7	94.8	96.9	95.9
60 to 64 years	96.1	94.1	95.4	95.3	96.0	94.2	95.0
65 to 69 years	91.9	97.2	98.2	98.4	92.4	97.3	95.1
70 years or more	99.3	100.4	99.6	97.9	99.3	100.2	99.8
15 years or more	91.4	94.5	76.7	87.4	89.8	92.7	91.8
All ages	91.8	94.9	79.3	86.4	90.2	93.1	92.1

¹Rates are given as percentages

SOURCE: Current Population Survey, October 1994: School Enrollment Technical Documentation.

Table 4. NHIS coverage rates^{1,2}, 3rd Quarter 1992

Age	Non-Black			Black		
	Male	Female	Both sexes	Male	Female	Both sexes
Under 1 year	85.1	94.5	89.7	92.1	81.1	86.6
1 to 4 years	92.2	95.2	93.6	76.5	95.1	85.7
5 to 9 years	93.2	96.5	94.8	92.0	78.0	85.2
10 to 14 years	94.3	99.6	96.9	83.7	82.9	83.3
15 to 17 years	96.7	102.4	99.5	73.8	85.8	79.7
18 to 19 years	94.3	88.3	91.3	53.9	64.8	59.5
20 to 24 years	88.8	87.3	88.1	66.0	78.9	72.9
25 to 29 years	88.6	85.3	87.0	63.0	90.6	78.1
30 to 34 years	85.2	91.2	88.2	81.6	99.3	91.3
35 to 44 years	89.0	94.4	91.7	73.1	91.0	83.0
45 to 49 years	87.5	97.4	92.5	85.8	82.4	83.9
50 to 54 years	89.3	91.0	90.2	58.2	86.6	74.1
55 to 64 years	N/A	N/A	N/A	74.3	92.9	84.5
65 to 74 years	96.9	95.9	96.3	66.6	83.4	76.2
75 years or more	92.1	88.9	90.1	90.1	92.6	91.7
All ages	91.0	93.4	92.2	75.7	87.1	81.8
All races	90.9					
Male	89.2					
Female	92.6					

¹Rates are given as percentages.

²These coverage rates do not take decennial census undercoverage into account.

SOURCE: Unpublished Census Bureau tabulations