METHODS TO IMPROVE THE PRECISION OF HEALTH STATISTICS FOR NON-HISPANIC ASIANS IN THE NATIONAL HEALTH INTERVIEW SURVEY

Carrie Jones, National Center for Health Statistics (NCHS)
Chris Moriarity, U.S. General Accounting Office
Karen Davis and Joe Fred Gonzalez, Jr., NCHS
Carrie Jones, NCHS, 6525 Belcrest Road, Room 915, Hyattsville, MD 20782

KEY WORDS: relative standard error, screening, density stratification

1. Introduction[1]

The National Health Interview Survey (NHIS) is a major national household survey conducted by the National Center for Health Statistics (NCHS) to monitor the health of the U.S. civilian noninstitutionalized population. The NHIS is redesigned following each decennial census to achieve new survey requirements and reflect changes in the distribution of the U.S. population. The current (in use since 1995) NHIS sample design oversamples Black and Hispanic persons; however, there is no special procedure to sample Asian persons, and current statistics for this group do not meet desired precision requirements for the next NHIS design. One of the primary research objectives for the next NHIS design is to improve the precision of annual estimates for Non-Hispanic (NH) Asian persons at the national level.

This paper discusses combining several years of NHIS data to make multi-year estimates and making estimates for collapsed age categories as two simple approaches to improve precision. Both of these methods do not require changes in the current sample design. This paper also presents findings on the effects of changing the current screening protocol to retain households containing Asians, and revising density substrata definitions with retaining all Asians as two alternative sample design changes.

2. Current Sample Design

The current NHIS is a stratified, multistage probability sample design. The four stages of sample selection are counties, including county equivalents, or groups of counties; clusters of housing units; households; and persons.

The NHIS oversamples Black and Hispanic persons by sampling geographical areas with high concentrations of Black and Hispanic populations at a higher rate than a national household survey of the same size where all households would have the same probability of selection. Another oversampling strategy includes household screening. That is, the field representatives attempt to administer a small subset of NHIS questions to obtain the race-ethnicity composition of households at some addresses. Households assigned to the screening interviews that do not contain any Black or Hispanic members are excluded from the sample and categorized as out of scope or "screened out"; all other households are retained in the sample and administered the full NHIS questionnaire.

3. Current Precision of Estimates

Before exploring methods for improving the precision of NH-Asian statistics, it is important to assess the current precision levels, in terms of the relative standard error (RSE), of a wide range of estimates. Fourteen health characteristics with low to high prevalence estimates (referred to as p) from the person, sample adult, and sample child sections of the NHIS and four analytical domains were selected for analysis. Table 1 contains a list of the health related characteristics. The domains are total, sex, age (<18 years, 18-44 years, 45-64 years, and 65+ years), and sex-age. SUDAAN [1] (Taylor series linearization approach) and SAS software were used to compute prevalence estimates and estimates of RSEs for the 210 (14 characteristics + 2 sexes x 14 characteristics + 4 ages x 14 characteristics + 2 sexes x 4 ages x 14 characteristics) health estimates.

For purposes of this research, it was determined that the NHIS would have the capability of producing estimates for NH-Asians at acceptable levels if the following precision criterion was met for each analytical domain: of the 210 estimates with $p \geq 10\%$, $RSE_1 \leq RSE_0$, where RSE_1 is the estimated RSE of the prevalence estimate computed from SAS/SUDAAN software and RSE_0 is the maximum acceptable RSE for each analytical domain (RSE_0=20% for total and sex domains, RSE_0=25% for age domain, and RSE_0=30% for sex-age domain).

[1] This paper does not necessarily reflect the views or position of the U.S. General Accounting Office.

Table 1: Health Related Characteristics

NHIS Section	Health Related Characteristics	
Person	Health insurance Activities of daily living	Hospital stay Health status
Sample Adult	Current smoker Ever smoked Diabetes Hearing Cancer	Hypertension Overweight Obesity Heart disease Asthma
Sample Child	Hearing	Asthma

Table 2 shows that the current NHIS does not have sufficient sample to provide acceptable annual NH-Asian statistics for all domains (denoted by cells with X) based on data from 1997 and 1998 separately. However, many cells do have sufficient sample.

Table 2: Domains Not Meeting The Precision Criterion for NH-Asian Persons

Domain	1997 NHIS	1998 NHIS
Total		
Sex		
Male		
Female		
Age		
<18		
18-44		
45-64		
65+	X	
Sex-Age		
Male, <18		
Male, 18-44		
Male, 45-64		
Male, 65+	X	X
Female, <18		
Female, 18-44		
Female, 45-64	X	X
Female, 65+	X	X

The amount of NH-Asian sample size increase to meet the precision criterion was estimated by solving for x in Formula 1:

$$RSE_0 = RSE_1 \cdot \sqrt{\frac{n}{n+x}} \qquad (1)$$

where RSE_0 is the maximum acceptable RSE for each analytical domain; RSE_1 is the estimated RSE of the annual prevalence estimate computed from SAS/SUDAAN software; and n is the current sample size of NH-Asians. The amount of sample size increase was equal to the largest value of x for the 210 prevalence estimates with $p \geq 10\%$ that do not satisfy the precision criterion.

There were 3,051 NH-Asian persons interviewed in the 1997 NHIS. An additional 5,327 NH-Asian persons (175% sample increase) would be necessary to produce statistics with acceptable precision. Based on the 1998 NHIS with 2,874 NH-Asian persons interviewed, the NHIS would require a 74% sample increase. Contributing factors to the different results include changes in prevalence levels from year to year; in particular, prevalence levels of a given health characteristic being ≥ 10% one year and < 10% the other year.

4. Methods For Improving Precision

Several constraints were adopted throughout the investigation of each method. First, the methods for improving precision for NH-Asian estimates must maintain the current precision for Hispanic and NH-Black estimates. Second, the overall NHIS sample size should remain unchanged at approximately 100,000 persons; thus, any increase in NH-Asian sample persons corresponds to an equal decrease in NH-"Other" (i.e., not Black or Asian) sample persons. The third constraint is the cost of the next NHIS sample should be similar to the cost of the current NHIS. Under some sampling scenarios, screening costs could increase, requiring a reduction of the overall sample size to keep the overall cost fixed. However, these cost considerations were not investigated.

4.1 Multi-Year Estimates

Although the primary redesign objective is to produce annual estimates for NH-Asians with acceptable precision, an alternative is to combine two years of NHIS data to make biennial estimates. A drawback to this method is precision estimates are slightly lower than they would be from independent samples because each year's sample is in the same set of primary sampling units. One must also assume that the prevalence rates and design effects do not change significantly between the two survey years.

This method was explored using combined 1997-1998 NHIS data. Prevalence estimates and estimates of RSEs were computed for the 210 health estimates. Based on the precision criterion in Section 3, the NHIS is capable of producing two-year estimates for NH-Asian persons at acceptable levels.

4.2 Annual Estimates For Collapsed Age Categories

Although the preferred age domain used in the research consisted of four categories, three age categories were also considered: <18 years, 18-44 years, and 45+ years. Prevalence estimates and estimates of RSEs were computed for the 168 health estimates using 1997 NHIS data. Based on the precision criterion in Section 3, the NHIS is capable of producing annual estimates for NH-Asian persons at acceptable levels for the collapsed age categories.

4.3 Changing Screening Protocol To Retain Households Containing Asians

In the current NHIS about 16,000 households are screened out of the sample every year. These are households that are assigned to be screened, but do not contain any Hispanic or Black persons. Another method considered for increasing the NH-Asian sample is to maintain the current sample design, but change the NHIS instrument to retain households containing Asian persons.

The estimated number of Asian persons in the screened out households was tabulated by separately accessing 1997 and 1998 NHIS unedited computer assisted personal interview (CAPI) data collected by the U.S. Bureau of the Census and applying a nonresponse adjustment. The estimated percentage increase in the NH-Asian sample was 77-80% (3,051 to 5,415 in 1997 and 2,874 to 5,161 in 1998), and the estimated percentage decrease in the NH-Other sample was between 3% and 4% to keep the overall NHIS sample size unchanged.

The estimated revised RSEs (RSE_2) of prevalence estimates for NH-Asian and NH-Other persons based on the revised sample sizes were computed using Formula 2:

$$RSE_2 = RSE_1 \cdot \sqrt{\frac{n}{r}} \qquad (2)$$

where RSE_1 is the estimated RSE of the annual prevalence estimates for the race-ethnicity group; n is the current sample size for the race-ethnicity group; and r is the revised sample size for the race-ethnicity group (r=n+i for NH-Asian and r=n-i for NH-Other, where i is the estimated number of screened out Asians).

The estimated RSE_2 of prevalence estimates for both race-ethnicity groups were evaluated by applying the precision criterion in Section 3 with the exception that there were a total of 294 health estimates for NH-Others versus the 210 for NH-Asians because 6 age categories were used. The precision criterion was not met for NH-Asian persons using 1997 data, but it was met using 1998 data. The one category lacking sufficient sample in 1997 was NH-Asian male persons aged 65 or older. On the other hand, the precision criterion was met for NH-Other persons for both survey years. Comparing these results with Table 2 suggests that changing the screening procedure to retain Asians does improve the precision of estimates for NH-Asian persons. The increased cost of retaining all Asians can be kept close to zero (i.e., cost-neutral) by increasing the screening rates for NH-Others, so that approximately the same number of additional NH-Others are screened out as additional NH-Asians are screened in.

This method is conceptually straightforward to implement. It can be made cost-neutral, has no negative impact on the precision of estimates for NH-Blacks and Hispanics, has little negative impact on the precision of estimates of NH-Others, and provides a substantial increase in the NH-Asian sample size. Thus, it is very likely that the next NHIS design will retain households with Asians.

4.4 Revising Density Substrata With Screening

Although retaining all Asians in the sample provides substantial improvement in the precision of estimates for NH-Asians, relative to the current NHIS design, the improved precision levels are still less than those obtained from the current NHIS for NH-Blacks and Hispanics. Thus, making density substrata definition changes in addition to retaining all Asians in sample were investigated as another method for providing further gains in the precision of estimates for NH-Asians. These two techniques are most effective when a large portion of the targeted population is highly concentrated in specific areas, as is the case with Black and Hispanic persons. Except for areas of high density in California and Hawaii, most Asians tend to live in more integrated neighborhoods which make them a difficult group to sample using these methods. Nevertheless, these methods may have some promise for Asians.

Given the selection in late 2001 of the PSUs for the next NHIS design, and the availability of 2000 Census data at the block level, it was decided to use the 2000 Census data in the NHIS PSUs to create a hypothetical universe from which simulated samples could be selected and analyzed. A sample allocation model with the revised density substrata was employed to obtain sample allocations that would be considered optimal, given the model assumptions, for the precision of NH-Asian estimates. This approach could have a negative impact on the precision of estimates for NH-Blacks and Hispanics; therefore, many different allocations of the sample across the revised density substrata were explored. The overall NHIS sample size was kept constant (as a proxy for cost neutral), and the loss of precision of estimates for NH-Blacks and Hispanics was not permitted.

The current sample design has 20 density substrata in the area frame defined by 1990 Census densities of Black and Hispanic populations [2]. The permit frame may be considered to be a separate density substratum. A revised set was developed and used in this research. It consisted of collapsing the set of 20 original density substrata into 6 substrata and then introducing a cross-classification of 10% "low-high" NH-Asian density which resulted in a total of 12 revised density substrata. Table 3 shows both the original and revised density substrata definitions. Note the original substrata definitions are density of Blacks but were changed to NH-Blacks for the revised substrata.

Table 3: Original and Revised Density Substrata

Revised Sub-stratum Label	Revised Density Substratum			
		Original Substratum		
	Percent NH-Asian	Label	Percent NH-Black	Percent Hispanic
1 2	[0,10) [10,100]	1	[0,10]	[0,5)
3 4	[0,10) [10,100]	6 11	[10,30) [30,60)	[0,5) [0,5)
5 6	[0,10) [10,100]	12 13 16 17 18	[30,60) [30,60) [60,100] [60,100] [60,100]	[5,10) [10,30) [0,5) [5,10) [10,30)
7 8	[0,10) [10,100]	2 3 7 8	[0,10) [0,10) [10,30) [10,30)	[5,10) [10,30) [5,10) [10,30)
9 10	[0,10) [10,100]	4 9 14 19 20	[0,10) [10,30) [30,60) [60,100] [60,100]	[30,60) [30,60) [30,60) [30,60) [60,100]
11 12	[0,10) [10,100]	5 10 15	[0,10) [10,30) [30,60)	[60,100] [60,100] [60,100]

Assuming equal unit variance in all strata, the sampling interval optimal for NH-Asians in substratum i is defined in Formula 3:

$$SI_i = k\sqrt{\frac{P_i}{D_{1i}}R - (R-1)(1-r_i)\frac{N_4 D_{4i}}{ND_{1i}}} \qquad (3)$$

where k is the constant of proportionality; R is the ratio of the cost of full interview to screened interviewed; N is the total population; N_x is the total population in domain x, where x takes on values of 1=NH-Asians, 2=NH-Blacks, 3=Hispanics, and 4=NH-Others; P_i is the population proportion in substratum i; r_i is the screening rate for NH-Others in substratum i; and D_{xi} is the proportion of domain x that falls in substratum i (i.e., the sum over i of D_{xi} is 1 for each x=1, 2, 3, 4, and the sum over x of $N_x \cdot D_{xi} = NP_i$ for each value of i). For sampling intervals optimal for NH-Blacks, substitute D_{2i} for D_{1i}; for Hispanics, substitute D_{3i} for D_{1i}.

Formula 3 was applied to Census blocks to obtain optimal sampling intervals for NH-Asians, NH-Blacks, and Hispanics in each of the revised density substrata using R=1.25. This value of R was obtained from cost modeling research conducted at NCHS [3]. Screening rates were computed in a way similar to the method used to obtain screening rates from the current NHIS design [4]; however, the newly-computed rates were based on the distribution of the 2000 Census data in the 12 density substrata, and assumed that all households containing Asians would be retained for interview. Table 4 contains the optimal sampling intervals for each race-ethnicity group.

The U.S. Bureau of the Census and NCHS determined that sample designs being considered should have costs roughly equivalent to a self-weighting design with 47,000 completed interviews. Assuming a nonresponse/noneligible rate of 20%, the target housing unit sample size becomes 47,000/.8=58,750. Using the 2000 Census count of 115,904,641 housing units, an appropriate self-weighting sampling interval to use is 115,904,641/58,750 ~ 1973.

Sets of 5 simulated samples were selected using the sampling intervals shown in Table 4 and the self-weighting sampling interval, for comparative purposes. Prevalence variables were simulated, variances of these simulated variables were estimated, and the results were averaged over the set of 5 simulated samples. As development took place to select the simulated samples, it quickly became apparent that many of the PSUs were too small to support a set of 12 density substrata. An algorithm was developed to combine density substrata, as necessary, by an automated process to assure that sample cases would be selected from all of the (combined, if necessary) substrata. In some PSUs, this led to all substrata being collapsed into one. When substrata were combined, the sampling interval from the density substratum that contributed the largest number of housing units was used for sampling.

Table 4: Optimal Sampling Intervals

Revised Substratum Label	NH-Asians	NH-Blacks	Hispanics
1	2987.5256	4830.1679	5271.9892
2	585.9975	3716.3004	3777.9626
3	2579.5100	972.8727	4427.1908
4	645.1459	1045.3864	3710.0476
5	3440.0519	595.0190	2283.0123
6	720.0495	757.8472	1531.9253
7	1990.5439	2241.4776	1425.9350
8	586.6109	2076.0626	1408.5297
9	2127.2747	1377.0621	798.2264
10	637.8176	1877.1915	816.6620
11	2806.5874	2395.1473	599.1646
12	763.9515	3086.8679	644.3939

The objective of the simulation study was to determine whether it might be possible to increase the effective sample size of NH-Asian households while maintaining the current target effective sample sizes of 6,000 NH-Black and 5,860 Hispanic households [4]. Table 5 shows the effective sample size of households for each sample allocation to the race-ethnicity groups. (Not shown are the effective sample sizes of households for the self-weighting design: All=46,776; Hispanic=3,858; NH-Black=4,651; NH-Asian=1,403; and NH-Other=36,995.) Not surprisingly, the effective sample sizes for NH-Asian households based on sampling intervals optimal for NH-Asians are larger than those obtained using a self-weighting sampling interval.

It can also be seen from Table 5 that the sampling intervals optimal for NH-Asian produces the largest effective sample size for NH-Asian households, but the effective sample sizes for NH-Black and Hispanic households are lower than targeted values. According to sampling intervals optimal for NH-Blacks, the effective sample size for NH-Black households exceeds the target of 6,000; however, the effective sample size for Hispanic households misses the target of 5,860. Likewise, using sampling intervals optimal for Hispanics, the effective sample size for Hispanic households exceeds the target of 5,860, whereas the effective sample size for NH-Black households misses the target of 6,000. Thus, the figures show that the sample allocation model defined in Formula 3 provides optimal allocation for one target domain, but it does not provide optimal allocation for the three target domains simultaneously.

Table 5: Effective Household Sample Sizes By Sampling Allocation

Race-Ethnicity Category	Sampling Allocation Optimal For		
	NH-Asians	NH-Blacks	Hispanics
All	31,243	23,696	25,045
Hispanic	4,128	3,743	8,421
NH-Black	4,019	9,669	4,759
NH-Asian	2,086	1,210	1,425
NH-Other	22,689	16,502	17,617

Note: Effective sample sizes need not be additive; i.e., the sum of the effective samples sizes across groups need not equal the effective sample size for the "All" category for a given sampling allocation.

Since the optimal sampling intervals cannot be used in practice because they do not meet one or more sample targets, linear combinations of these sampling intervals were developed. In addition, a new redesign requirement was imposed in early 2002 to limit the variability in the base weights resulting from the first two stages of sampling. As can be seen in Table 4, using optimal sampling intervals for NH-Asians introduces a base weight variability factor as high as 3,440/586 ~ 6. For these reasons, simulated samples were selected using sampling intervals equal to linear combinations of the optimal sampling intervals and a self-weighted interval, pulled toward the self-weighting sampling interval (no weight variability) until the base weight variability factor of at most 2 was met.

A large number of compromise sampling intervals were developed. Three of the five that met the base weight variability requirement are listed below. Compromise sampling interval A denotes 0.3 times the optimal sampling interval for NH-Blacks plus 0.7 times the self-weighting sampling interval. The other compromise sampling intervals are defined accordingly.

(A) 0.3 SI_NHB + 0.7 SW
(B) 0.3 SI_HIS + 0.7 SW
(C) 0.18 SI_HIS + 0.12 SI_NHB + 0.7 SW

Table 6 shows the effective sample size of households for each compromise sample allocation to the race-ethnicity groups. Although each set of sampling intervals met the weight variability criterion, the effective sample sizes of NH-Black and Hispanic households were fewer than those in the current design, except for the 6,441 NH-Black households using the set of compromise sampling intervals defined by A.

Table 6: Effective Household Sample Sizes By Compromise Sampling Allocation

Race-Ethnicity Category	Compromise Sampling Allocation		
	A	B	C
All	32,738	31,125	31,981
Hispanic	4,397	4,945	4,773
NH-Black	6,441	5,351	5,773
NH-Asian	1,424	1,534	1,445
NH-Other	23,175	22,012	22,669

The results of the simulation study suggest that it is not possible to simultaneously obtain a large gain in the effective sample size for NH-Asian households while maintaining the current target effective sample size for NH-Black and Hispanic households, satisfying the weight variability restriction criterion, and keeping a cost-neutral design. More interesting, it does not appear to even be possible to maintain the current effective target sample size for NH-Black and Hispanic households regardless of the effective sample size of NH-Asian households. These findings are not surprising for several reasons. First, the base weight variability restriction criterion is more stringent than the current design allowance of ~ 2.3. Second, the estimated ratio of interviewed to screened cost is 1.25 as compared to 3 for the current design [4]. Third, a 90% response rate, which is a lower rate than 10 years ago, is postulated.

It is unlikely that other substrata definitions would give results that are much different than those presented in here; thus, if the weight variation restriction is required and the ratio of interview to screening cost assumption is accurate, it is reasonable to conclude that a cost-neutral budget for the next redesign may not be able to meet current target precision levels for NH-Blacks and Hispanics. Nevertheless, it is recommended that substrata definitions and sample allocations that have similarities to those described here be adopted for the next NHIS design.

In many respects, the simulation results presented here were comparable to findings from a model-based line of sample allocation research conducted independently at NCHS that will most likely be used to specify the within-PSU design parameters for the upcoming design.

5. Conclusion

The results presented in this paper demonstrate that the following methods could provide improved precision of NHIS estimates for NH-Asian persons:

- Multi-Year Estimates
- Annual Estimates For Collapsed Age Categories
- Changing Screening Protocol To Retain Asians
- Revising Density Substrata With Screening

A final decision has not been made as to which of these methods, if any, will be implemented in the next NHIS design, but it is very likely that the current screening protocol will be changed to retain households containing Asians and a model-based line of sample allocation research will be used to specify within-PSU design parameters.

References

1. Shah, B., Barnwell, B., Bieler, G. (1997). *SUDAAN Users's Manual, Release 7.5*, Research Triangle Park, NC: Research Triangle Institute.

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

3. Shimizu, I., Lan F. (2001). *Approximation of Variable Costs for the National Health Interview Survey*. 2001 Proceedings of the American Statistical Association, Section on Survey Research Methods [CD-ROM]. Alexandria, VA: American Statistical Association.

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