DESIGN EFFECTS FOR THE NATIONAL HOSPITAL AMBULATORY MEDICAL CARE SURVEY

Iris Shimizu, Kenrick Henderson, and Wayne Johnson, National Center for Health Statistics
Iris Shimizu, National Center for Health Statistics, 6525 Belcrest Rd., Room 915, Hyattsville, Maryland 20782

Key words: Sampling errors, establishment survey

1. Introduction[1]

Design effects for standard errors from a sample are ratios in which the numerators are the standard errors actually realized in that sample and the denominators are the standard errors that would have been calculated if the sample were a simple random sample. (Design effects for variances are similarly defined but are calculated with variances in place of the standard errors.)

Design effects for complex samples can be useful. For example, they are used to determine the effective sample sizes of complex samples, where effective sample size is the size of a simple random sample that would produce precision levels equal to what was obtained in the complex sample. Design effects may also be used in adjustments of test statistics such as chi-square statistics to account for complex sample designs. Design effects can also be used to approximate the sampling variances of statistics from complex surveys when one does not have the means to correctly compute the variances, either because the required sampling design information is lacking or because the computer software which correctly approximates variances for complex sample statistics is unavailable.

This paper describes research undertaken to develop design effects for several classes of variables in the National Hospital Ambulatory Medical Care Survey (NHAMCS) which the National Center for Health Statistics conducts to produce statistics about the numbers and kinds of visits to hospital emergency and outpatient departments. The NHAMCS uses a stratified four-stage probability sample. However, the public cannot be given some of the design information required for variance computations because that information poses a risk to the confidentiality of respondents. Design effects offer an alternate way to approximate the NHAMCS variances.

The next section describes the NHAMCS sampling design and estimation procedures used to approximate the variances of estimates. Section 3 discusses the methodology used to derive design effects and section 4 presents the results. Section 5 gives an illustration of how to use the design effects to approximate variances for NHAMCS statistics. Section 6 summarizes the results of the research.

[1] The opinions expressed in this paper are those of the authors and not necessarily those of the National Center for Health Statistics

2. Survey design

Sample

The NHAMCS universe consists of in-person visits to emergency departments (EDs) and outpatient departments (OPDs) of non-Federal, non-institutional short stay hospitals (average length of inpatient stay is less than 30 days) or general medical or general surgical hospitals (without regard to length of stay) which had at least six beds set up and staffed for inpatient care. The sampling frame consists of hospitals in SMG's 1991 Hospital Market Database (SMG 1991) which satisfies the criteria for being in the survey universe. The first stage sample consists of 112 primary sampling units (PSUs) which is a probability subsample of the PSU's selected to the 1985-94 National Health Interview Survey. The PSUs are counties (county equivalents) or groups of counties, except in New England where some PSUs are formed from townships. The PSUs are stratified by region, socioeconomic, and demographic characteristics and one PSU was selected from each stratum. The second stage sample consists of 600 hospitals which are divided into 16 national sample panels which were randomly ordered for assignment to 4-week reporting periods. Each year, data collection is attempted at the approximately 490 hospitals assigned to the 13 reporting periods falling within that year. The third stage sample is a stratified sample of service areas where strata are defined by department within the sampled hospital. Emergency service areas are selected from the emergency department or ED and outpatient clinics are selected from the outpatient department or OPD. In departments that have more than five service areas, a sample of five service areas is selected without replacement and with probability proportional to size where size is the number of visits expected in the clinic during the hospital's assigned reporting period. If the department has five or fewer outpatient clinics, then all of that department's service areas are included in the sample. The fourth stage sample consists of systematic random samples of visits to the selected service areas.

Design effects were calculated for estimates produced from the 1995-1996 NHAMCS sample data sets. In 1996, the sample included 486 hospitals of which 438 were NHAMCS- eligible (that is, had either an ED or an OPD) in 1996. Of these 438 hospitals, 95 percent participated in NHAMCS with 392 EDs providing 21,902 patient visit abstracts and 235 OPDs providing 29,806 patient visit abstracts. In 1995, the sample included 487 hospitals of which 437 were NHAMCS-eligible in 1995. Of these 437 hospitals, 94 percent participated in NHAMCS with 391 EDs providing 21,911 patient visit abstracts and 230 OPDs providing 28,393 patient visit abstracts.

Estimation

Statistics from NHAMCS are derived by a multistage estimation procedure that produces essentially unbiased estimates. The estimation procedure has three basic components: (a) inflation by reciprocals of the sampling selection probabilities, (b) adjustment for nonresponse, and (c) a population weighting ratio adjustment.

Inflation by reciprocals of selection probabilities

There is one probability for each sampling stage: (a) the probability of selecting the PSU, (b) the probability of selecting the hospital, and (c) the probability of selecting the service area within the department, and (d) the probability of selecting the visit from the year within the service area. The last probability is calculated to be the sample size from the service area divided by the product of 13 (number of 4-week data collection periods in a year) times the total number of visits that actually occurred in that unit during the hospital's assigned data collection period. The overall probability of selection is the product of the probabilities at each stage. The inverse of the overall selection probability is the basic inflation weight.

Adjustment for nonresponse

NHAMCS data are adjusted to account for two types of nonresponse. The first type of nonresponse occurred when a sample hospital refused to provide information about their EDs and/or OPDs which were publicly known to exist. In this case, the weights of visits to hospitals similar to the nonrespondent hospitals were inflated to account for visits represented by the nonrespondent hospitals where hospitals were judged to be similar if they were in the same region, ownership control group (government, voluntary nonprofit, or proprietary), and metropolitan statistical area (MSA) status group (in an MSA versus not in an MSA) where MSA is defined by the Census Bureau. This adjustment was made separately by department type.

The second type of nonresponse occurred when a sampled service area within a "respondent" hospital failed to provide completed patient abstract forms for a sample of their patient visits. The weights for service areas that were similar to a nonrespondent service area were inflated to account for that nonrespondent service area where service areas were judged to be similar if their hospitals were in the same region, ownership control group, and MSA status group and if the services areas came from the same type of department (ED or OPD). For OPD service areas, similarity also meant that the service areas came from the same service group where service groups were: general medicine, pediatrics, surgery, OB/GYN, alcohol and/or substance abuse, and other specialty.

Ratio Adjustment

Adjustments were made within hospital strata defined by region, and within the South and West, the adjustment strata for EDs were further defined by hospital ownership groups. These adjustments were made separately for emergency and outpatient departments. For EDs, the adjustment was a multiplicative factor that had as its numerator the sum of annual visit volumes reported to EDs in sampling frame hospitals in the stratum and as its denominator the weighted sum of those visits for sample hospitals in that stratum. The adjustment for visits to OPDs was a multiplicative factor which had as its numerator the number of OPDs reported in sampling frame hospitals in the stratum and as its denominator the weighted number of those OPDs based on sample hospitals in that stratum. The data for the numerator and denominator of both adjustments were based on data recorded for hospitals in the April release of the SMG Hospital Market Data Base in the year following the reference year. For example, data in SMG's 1997 April release were used for the ratio adjustment in the 1996 NHAMCS statistics.

Variances

To compute the standard errors for NHAMCS statistics, NCHS uses the linearized Taylor Series approximation applied in the SUDAAN software (Shah 1997). Because only one PSU was selected from each stratum of non-self-representing PSUs, the non-certainty PSU strata were collapsed to permit the computation of PSU variances. Each collapsed stratum contained two or three sample PSUs.

To simplify variance computations, the non-ratio adjusted or preliminary weights (which include only the basic inflation weights and adjustments for nonresponse) are used with SUDAAN to produce variance approximations which are theoretically conservative (ratio adjustments theoretically reduce variances if the ratio denominator is correlated with the variable of interest). The ratio adjustment is calculated with hospital data and weights instead of visit data and weights. Because SUDAAN does not accommodate the simultaneous use of two weights (one for visits and one for hospitals), SUDAAN cannot calculate variances for NHAMCS in a single run.

3. Methods

The public use files containing NHAMCS data exclude design variables which pose risks to the respondents' confidentiality. Hence, the public data users are unable to adequately approximate sampling errors for the NHAMCS statistics. To assist users, we compare standard errors produced in SAS 6.12 with those produced in SUDAAN. Standard errors are computed for percentages and totals for several variables and domains using SUDAAN and three other approximate methods which we implemented within SAS. We calculate ratios of standard errors in which the numerator is produced by SUDAAN and the denominators are produced by the other approximate methods. For discussion purposes, these ratios of standard errors will be referred to as "design effects" (DEFFs).

All calculations are based on the "in-house" data sets which contain the NHAMCS survey design variables that are required to correctly calculate variances. The sampling unit is a patient visit for health care to either hospital outpatient services or emergency rooms. Each visit record in the "in-house" data file contains two versions of the sampling weight; the first is the preliminary weight (described in the prior section) and the second is the first weight multiplied by the ratio adjustment to minimize mean square errors. The ratio-adjusted (final) weight is the only weight included in public-use files and, hence, the only weight available to the public data user. The ratio-adjusted weight was used to produce all point estimates and all standard errors computed in SAS while the preliminary weight was used for standard errors in SUDAAN.

The user we had in mind is one that would be likely to lean upon available software for support. Hence, three methods of variance estimation which we thought worth including in the study are: (1) treating the data as though they were from a simple random sample, (2) using the weights in SAS's variance formula with the default denominator n-1 (VARDEF=DF), and (3) treating the data as if sampling with unequal probabilities. The first two of these techniques naturally estimate a proportion and are extended to totals by multiplying with a sample estimate of the population total; the third technique naturally estimates totals, so proportions are estimated by dividing by the estimated population total. The finite population correction factor is ignored in all three of the methods because the visit sample is small relative to the total population. Table 1 gives the formulas used in SAS to calculate the three alternate sampling error approximations for totals and percentages. Design effects calculated by using the first method (that is, by assuming a simple random sample) to compute standard errors are the "traditional" design effects.

A set of DEFFs was produced for each of six sets of NHAMCS statistics which users tend to analyze. These sets are annual estimates for visits to EDs, to OPDs, and to EDs/OPDs combined and biennial estimates for visits to EDs, to OPDs, and to EDs/OPDs combined. In each set, design effects were produced for estimates corresponding to those published in NCHS's Advanced Data reports about 1996 NHAMCS statistics for visits to EDs and OPDs (McCaig 1997; McCain and Stussman 1997).

4. Results

To conserve space, the results based on only 234 distinct estimates are summarized here to illustrate relationships that appear to be typical of those observed among the NHAMCS design effects (DEFFs). These 234 estimates are a subset of the biennial estimates for visits to OPDs in 1995-96 and exclude estimates about specific drugs prescribed during the visits and estimates based only on the subpopulation of injury visits.

Table 2 displays design effects (DEFFs) for OPD visit estimates by selected demographic characteristics of OPD clients. The first three columns in the table include information that was published in McCaig (1997). Columns 4-6 present the design effects produced by the three different methods for the estimated numbers of visits while the last three columns give the design effects for the corresponding estimated proportions of visits. It can be seen that the DEFFs resulting from the first two methods for calculating DEFFs appear similar to each other. This similarity among the DEFFs from those two methods appears true for most of the statistics and domains included in the study. Because of the similar behavior among the first two DEFF methods, subsequent discussion may mention only the first method with the understanding that results from the first method are similar to those for the second method.

Other relationships between the DEFFs from the different methods also appear to hold for most statistics included in the study. In particular, the DEFFs from the three different methods are correlated with each other. For the set of 234 biennial variables for visits to OPDs, the correlations between the DEFFs from the three methods range between 0.94 and 1.0 for the aggregate estimates and from 0.90 to 1.0 for the percent estimates. Also, the DEFFs produced by the third method appear to be consistently smaller than those from the other two methods. For the set of 234 biennial aggregate estimates for OPD visits, the means of the DEFFs from the first two methods are 11.2 and 11.1, respectively, while the smaller mean of the DEFFs from the third method is 5.3. For the corresponding percent statistics, the means of the DEFFs from the first two methods are 5.9 and 5.9 as compared to the smaller mean of 3.0 for the DEFFs from the third method.

As can frequently be expected, the DEFFs produced for NHAMCS estimated numbers of visits tended to be correlated with the magnitude of the estimate. For the set of 234 estimates, the correlations between the estimated visit counts and the corresponding DEFFs from the different methods ranged from 0.89 to 0.93. As can be expected, the DEFFs for estimates of visits by some populations that are probably clustered appear to be greater than the DEFFs for comparable sized estimates for populations that are not clustered. For example, note in Table 2 that approximately 8.8 and 8.9 million visits were made to OPDs by Black children under the age of 15 years and by Blacks aged 25-44 years, respectively. While the estimated numbers of visits are similar for the two groups, the three DEFFs for the estimated number of childrens' visits range from 5.3 to 11.2 while the corresponding DEFFS for the 25-44 year olds appear lower, ranging from 3.7 to 5.9. It is possible that the Black children tended to cluster in clinics that specialized in pediatric care while the visits by the group of older Blacks were dispersed across a greater variety of OPD clinics.

To examine the extent of variation in DEFF values for NHAMCS, it is sufficient to examine the distribution of DEFF values calculated by the first or traditional method. For the set of 234 biennial estimates about visits to OPDs, the following values were observed for the DEFFs:

Statistic type	Aggregate	Percent
mean	11.2	5.9
SE	10.0	3.7
Median	8.2	4.9
Range	0.9-55.3	0.9-30.8

The DEFFs for some OPD statistics could be expected to have considerable size. First, the between-hospital variance component for OPD statistics was increased by the existence of NHAMCS-eligible sample hospitals which had an ED but not an OPD. Also, when service areas (OPD clinics) were sampled, the between-clinic (within hospital OPD) variance component was affected by the specialization that usually existed between clinics in those OPDs which had multiple clinics; that is, some visit and patient characteristics were likely to be clustered by clinic specialty. While not presented in this paper the DEFFs observed for ED statistics tend to be less that those for OPD statistics, which would be expected because service areas in EDs are rarely sampled. Also, hospitals are more likely to have EDs than OPDs with the result that the between-hospital variance components should also be less for ED statistics than for OPD statistics.

5. Illustration of DEFF use

To illustrate the use of DEFFs to approximate variances of NHAMCS estimates, let us assume that the user has combined the NHAMCS Public Use files for 1995 and 1996 and he wants an approximation for the standard error of the percent of OPD visits made by females in 1995-6. Table 2 shows that an estimated 61.5 percent of the visits to OPDs in those years were made by females. The 61.5 percent is a weighted mean for the binomial variate defined by "x = 1 if patient is female and x= 0 otherwise." According to the formulas given in Table 1, this weighted mean can be used in the method two formula for calculating standard errors when one has a binomial variable. Using that formula and the combined 1995-96 OPD sample size of 58,199 (from Section 2, above) gives a standard error estimate of

$$\sqrt{\frac{\hat{\mu}(1-\hat{\mu})}{n-1}} = \sqrt{\frac{0.615(1-0.615)}{58,198}} = 0.0020 \ . \qquad (1)$$

Table 2 shows that, when the second method is used to calculate standard errors, the corresponding DEFF for the standard error of the estimated percent of visits by females in 1995-96 is 3.4. Multiplying the result in (1) by the corresponding DEFF gives $3.4 \times 0.0020 = 0.0069$. The standard error for the NHAMCS estimate of 61.5 percent of visits by females in 1995-96 is thus about 0.7 percentage points.

6. Summary

NHAMCS uses a complex sample. However, design variables required to correctly calculate variances pose risks to the confidentiality of survey respondents. Thus, those variables are excluded from the public use data files and users of those files are unable to compute their own "exact" variances. Design effects (DEFFs) would enable the data users to adequately approximate those variances.

This paper describes three methods for calculating design effects for standard errors. From the three sets of DEFFs produced based on the different methods, a user would choose the set of DEFFs that corresponds to the method which the user chooses for calculating variances from the NHAMCS public use files.

This paper also presents DEFFs produced by using the three DEFF calculation methods for biennial statistics for visits to hospital outpatient departments and examines their properties. The DEFFs resulting from two of the three methods are similar for most statistics, and the

DEFFs produced under all three methods were correlated with each other. The DEFFs appear to vary considerably with no clear pattern on which to base guidance about the best choice of DEFF values to use when approximating the sampling errors for NHAMCS estimates. The DEFFs for aggregate statistics do appear to be correlated with the magnitude of the estimates. Thus, it appears that choice of the DEFF value to use in approximating errors for at least aggregate statistics should probably depend upon the magnitude of the statistic. The choice of DEFF value for use in approximating errors for percent statistics is not so clear.

In addition to DEFFs for biennial statistics for OPD visits, DEFFs are also being calculated for annual estimates for OPD visits and for both annual and biennial statistics about visits to EDs and to EDs/OPDs combined. DEFFs for estimates about injury visits and drugs prescribed or given during the visits are also being produced.

References

McCaig, LF (1997). "National Hospital Ambulatory Medical Care Survey: 1996 Outpatient Department Summary." Advance data from vital and health statistics; no 293. Hyattsville, Maryland: National Center for Health Statistics.

McCaig, LF; and Stussman, BJ (1997). "National Hospital Ambulatory Medical Care Survey: 1996 Emergency Department Summary." Advance data from vital and health statistics; no 294. Hyattsville, Maryland: National Center for Health Statistics.

Shah, BV; Barnwell, BG; Bieler, GS (1997). SUDAAN User's Manual, Release 7.5. Research Triangle Park, North Carolina; Research Triangle Institute.

SMG Marketing Group Inc. (1991). 1991 Hospital Market Data Base. 1342 N. LaSalle Drive, Chicago, Illinois 60610. SMG Marketing Group Inc.

SMG Marketing Group Inc. (1997). Hospital Market Database, April 1997 Release. John Hancock Center, Chicago, Illinois 60611. SMG Marketing Group Inc.

Table 1: Parameters estimated and estimation formulas by method for calculating alternate sampling error approximations

	Parameter estimated			
Approximation Method	Mean $\hat{\mu}$	S.E. $\hat{\sigma}_{\hat{\mu}}$	Total $\hat{t}$	S.E. $\hat{\sigma}_{t}$
1. SRS	$\frac{1}{n}\sum_{i=1}^{n} x_i$	$\sqrt{\frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{n\cdot(n-1)}}$	$\hat{N}\cdot\hat{\mu}$	$\hat{N}\cdot\hat{\sigma}_{t}$
		$=\frac{\hat{\mu}(1-\hat{\mu})}{(n-1)}$ if x_i is binomial		
2. SAS (VARDEF =DF)	$\frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$	$\sqrt{\frac{\sum_{i=1}^{n} w_i(x_i-\hat{\mu})^2}{\sum_{i=1}^{n} w_i(n-1)}}$	$\hat{N}\cdot\hat{\mu}$	$\hat{N}\cdot\hat{\sigma}_{t}$
		$=\frac{\hat{\mu}(1-\hat{\mu})}{(n-1)}$ if x_i is binomial		
3. UNEQWR	$\frac{1}{\hat{N}}\cdot\hat{t}$	$\frac{1}{\hat{N}}\cdot\hat{\sigma}_{t}$	$\frac{1}{n}\sum_{i=1}^{n} n w_i x_i$	$\sqrt{\frac{\sum_{i=1}^{n}(n w_i x_i-\hat{\mu})^2}{n\cdot(n-1)}}$

where $\hat{N} = \sum_{i=1}^{n} w_i$.

Table 2: Design effects for estimated number and percent distribution of hospital outpatient department visits by selected patient characteristics in United States; 1995-96 NHAMCS.

	Number (in 1,000s)	Percent	Design effect method 1	2	3	1	2	3
			Design effects for numbers			Design effects for percents		
All visits	134418	100.0	.	.	17.0	.	.	0.0
Age								
Under 15 Years	30236	22.5	17.0	16.5	7.6	8.8	8.5	3.9
15-24 years	16617	12.4	10.2	10.4	5.4	3.8	3.8	2.0
25-44 years	37135	27.6	16.3	16.6	8.4	5.9	6.0	3.1
45-64 years	29722	22.1	14.8	14.6	7.4	3.9	3.9	2.0
65-74 years	11804	8.8	9.8	9.7	5.2	4.3	4.2	2.3
75 years and over	8904	6.6	13.3	13.4	6.9	8.9	8.9	4.6
Sex and age								
Females	82309	61.2	33.9	33.8	12.6	3.4	3.4	1.3
Under 15 years	14130	10.5	10.7	10.3	5.0	5.4	5.3	2.5
15-24 years	12271	9.1	9.0	9.3	4.9	4.2	4.4	2.3
25-44 years	24565	18.3	12.6	12.9	6.9	5.0	5.1	2.7
45-64 years	18605	13.8	11.7	11.5	6.0	3.0	3.0	1.6
65-74 years	7089	5.3	7.7	7.6	4.1	3.1	3.1	1.6
75 years and over	5649	4.2	11.6	11.6	6.0	8.3	8.3	4.3
Males	52109	38.8	20.9	20.8	9.3	3.4	3.4	1.5
Under 15 years	16105	12.0	12.1	11.7	5.6	6.5	6.3	3.0
15-24 years	4346	3.2	5.2	5.3	2.9	2.1	2.1	1.1
25-44 years	12570	9.4	8.6	8.8	4.8	3.3	3.4	1.9
45-64 years	11117	8.3	8.0	8.0	4.3	3.1	3.1	1.7
65-74 years	4714	3.5	6.4	6.3	3.5	3.8	3.8	2.1
75 years and over	3255	2.4	6.7	6.9	3.6	4.0	4.0	2.1
Race and age								
White	96755	72.0	50.5	50.9	16.1	12.0	12.1	3.8
Under 15 years	20520	15.3	15.6	15.1	7.0	7.7	7.5	3.5
15-24 years	11303	8.4	9.6	9.8	5.1	3.8	3.9	2.1
25-44 years	26786	19.9	15.7	15.9	8.2	6.0	6.1	3.2
45-64 years	21638	16.1	14.8	14.5	7.3	4.8	4.6	2.3
65-74 years	9153	6.8	9.8	9.7	5.2	4.7	4.7	2.5
75 years and over	7354	5.5	13.9	13.9	7.1	9.7	9.7	5.0
Black	33399	24.8	13.6	13.5	6.9	11.8	11.7	5.9
Under 15 years	8845	6.6	11.2	10.6	5.3	9.9	9.3	4.7
15-24 years	4794	3.6	5.7	5.8	3.1	4.6	4.6	2.5
25-44 years	8866	6.6	5.9	6.2	3.7	5.7	5.9	3.5
45-64 years	7094	5.3	5.9	5.9	3.4	5.6	5.6	3.2
65-74 years	2384	1.8	4.8	4.7	2.6	4.2	4.1	2.3
75 years and over	1416	1.1	4.2	4.2	2.3	3.8	3.7	2.0
Asian/Pacific Islander	3919	2.9	7.6	8.7	5.5	6.7	7.8	4.9
American Indian/	345	0.3	2.9	3.1	2.3	2.8	3.0	2.2
Geographic region								
Northeast	37596	28.0	20.0	20.7	10.8	18.7	19.4	10.1
Midwest	49692	37.0	51.1	47.8	17.9	30.8	28.8	10.8
South	31342	23.3	16.8	16.9	9.4	17.5	17.6	9.8
West	15787	11.7	13.7	15.7	13.3	13.8	15.9	13.4