A Comparison of the Effectiveness of Two Nursing Home Sample Designs

John Sommers, Agency for Health Care Policy and Research
2101 East Jefferson Street, Suite 500, Rockville, Maryland 20852

Key Words: Sample Design, Nursing Homes

Background

In 1987, The National Center for Health Services Research, the predecessor to The Agency for Health Care Policy and Research (AHCPR) conducted the National Medical Expenditure Survey - Institutional Population Component (IPC). This was a survey that included Nursing Homes (NH), Board and Care Homes (BC) and Facilities for Persons with Mental Retardation (FMR). The survey collected information on facility characteristics as well as information for a sample of patients. This information included items, such as, expenditures, demographics, and health status of a sample of persons staying within each facility. The sample of persons consisted of two parts, a sample of persons in the facility as of January 1, 1987 and a sample of persons admitted during 1987.

In 1996, AHCPR conducted The Medical Expenditure Panel Survey - Nursing Home Component (NHC). The survey was very similar to the IPC in content, but was restricted to nursing homes only.

The sample design for the 1996 survey was similar to that of the 1987 survey. Revisions were made after analysis of the 1987 data and other available information. (See Cohen, et. al., 1993 and Bethel, et. al., 1997) Many of these changes were made primarily to improve estimates of expenditures by persons in the NH's. However, it was felt that other types of variables would also benefit considerably from these changes. The purpose of this paper is to compare the precision of the estimates for the same sets of characteristics of nursing homes and their patients made from the two surveys in order to determine if the changes made to the sample design met our expectations by producing improved estimates and to what degree.

A Comparison of the Two Sample Designs

Both the IPC and NHC sample were multiple stage stratified samples, where the first stage sampling units were facilities selected from a list frame. The frame used for the IPC was the 1986 Inventory of Long Term Care Places (ILTCP), a list complied by the AHCPR, National Center for Health Statistics (NCHS) and Health Care Financing Administration (HCFA). This was a list of NH's, BC's and FMR's developed through use of past lists and information supplied by each state concerning licensed or known facilities within the state. (See Scirroco, 1989) Besides address information, the ILTCP contained information on facility type, certification, size and ownership which could be used for stratification.

The NHC was selected from the 1995 Health Provider Inventory (HPI), a list of potential NH's.
The 1995 HPI was produced from the 1991 HPI. The 1991 HPI was a list with similar facilities to that of the 1986 ILTCP. This list was updated by NCHS and AHCPR using information collected from the states by NCHS, plus information on hospital based NH's from the American Hospital Association and the Veterans Administration collected by AHCPR. (See Bethel, et. al., 1998, Scirroco, 1994 and Strahan, 1997) Because of problems with duplicate sampling units selected in the IPC, extra effort was taken by both NCHS and AHCPR to unduplicate the 1995 HPI to avoid this sampling inefficiency.

Before sample selection, each frame was divided into strata and within each stratum, each frame was also sorted by implicit strata. For the IPC there were three strata, (1) NH's and BC's, (2) small FMR's and (3) large FMR's. Within each stratum, the file was sorted by region, facility certification status, ownership type, number of beds, number of admissions, state and Zip code.

For the NHC, there were 7 primary strata, the first contained the 20 largest facilities. The other were based upon a cross of whether a facility was associated with a hospital and which of three Medicaid reimbursement methods was used within the state. Within each stratum the facilities were sorted by a location variable, ownership type and zip code. The location variable indicated whether a facility was located in a central city, suburb, adjacent to a metropolitan area or not adjacent.

The stratum of the largest facilities was chosen to allow AHCPR to share facilities that would have normally have been certainty facilities in the 1995 National Nursing Home Survey conducted by NCHS and/or the NHC. To coordinate sample and limit individual respondent burden, the two agencies agreed that no NH would be in both samples. A special Keyfitz selection process guaranteed this for all other facilities except these twenty. As a result the agencies agreed to share these facilities among the two surveys. In order to do this the facilities were placed in a separate stratum and one process was used to select which of the facilities was selected for the two individual surveys. The remaining explicit and implicit strata were determined after analysis of 1987 data. (See Bethel, et. al., 1993). The variables were selected primarily due to their correlation with expenditure variables. However, they also have some relation to many other variables collected. For instance, hospital based NH's tend to offer more expensive rehabilitative type care. But their clientele also tend to be younger and have somewhat different health characteristics, so it is a useful in stratification for these variables also. Likewise, the location variables are very useful for expenditures, but also relate to patient characteristics, such as race and income.

For each survey, a first phase sample of facilities was selected using a systematic selection process with probability proportional to a total measure of size within the stratum. The measures of size used differed across the two surveys. For the 1987 survey, the relative measure of size was the largest of two proportions. The first was the proportion of beds in the facility of all beds in the stratum. The second was the proportion of admissions in the facility compared to the total admissions for all facilities within the stratum. Both variables were reported on the frame for a previous time period. The logic for this selection of measure of size was that the survey was interested in both current residents and admissions and thus the measure of size was the most important of the two.

The 1996 NHC used number of beds as the measure of size. This choice was made after analysis of results from the 1987 survey. It was found that the number of residents in a facility and the numbers of admissions during the year both had a higher correlation with the number of beds reported on the frame than with any other available variable on the frame. (Bethel, et. al., 1993)

For each survey the first sample selected was divided into four groups, dependent upon their distance from available interviewers and the relative travel costs required to collect the data from the facility and subsampled at differing rates. The size of the subsample and the corresponding sampling weights were determined as follows. For the 1987 sample, using an assumption of equal variances within each stratum, a Neyman allocation (Cochran, 1977) was made using travel costs and a fixed travel budget. This resulted in different subsampling rates within each of the cost strata. Due to the variation in travel costs, subsampling rates ranged from a high of 96% to a low of 30%. This resulted in an approximate 10% increase in standard error compared to a proportional allocation of a sample of the same size. Under an assumption of equal strata variances, a proportional allocation of sample would be optimal if there were no cost constraints.

For the 1996 NHC, the first sample was also divided into four cost strata based the same criteria. However, rather than calculate an optimal Neyman allocation based upon travel costs alone and a fixed travel budget, the optimal allocation was determined using total collection estimated costs for facilities within each strata and a fixed total collection budget. This was done because AHCPR realized from the 1987 experience that travel was a much smaller percentage of the collection budget and total costs were the appropriate cost measure that should be used. This resulted in much smaller relative cost differentials across the four cost strata. Subsampling rates thus had a much smaller relative difference. The high was 89% and the low 78%. The estimated effect on the standard error compared to a proportional sample of the same size was about 1%.

During the field data collection for each survey, a sample of current residents, persons who resided in a nursing home on January 1, 1996, and admissions, persons whose first stay in a nursing home in 1996 started in 1996, were selected in each responding facility.

The expected sample size for each group for each survey was 4 per facility. This was done so that estimates could be made for person/patient level information, such as, health status of patients in NH's.

Again the two surveys differed slightly in their approach. To select current residents in both surveys a list of residents in the NH at midnight December 31 of the previous year was developed and a sampling rate systematically applied. However, the means used to calculate the sampling rate was different. In the 1987 survey, a sampling rate was predetermined for each NH so that the final probability of selection of each sampled current resident was identical. This rule assumed that the measure of size was an adequate representation of the actual number of current residents in the facility. The exception to this rule was that if the size of the list was very different than that predicted from the facility measure of size, then the sampling rate was changed so that the final number of current residents selected was always greater than 1 and less than 13. (Cohen et. al., 1993)

The 1996 survey, for the most part, fixed the sample size 4 within the facility. The rate was determined using frame information on the number of beds. This was done because research showed the new measure of size would guarantee that a fixed sample would yield almost equal weights on the current residents while greatly simplifying operations. However, if a sample of 4 created a weight per person that differed beyond a set limit from the expected value, the sample could be increased to 6 persons in order to control for excess weights.

These same differences were also reflected in the selection of admissions for the two surveys. The 1987 used a fixed probability of selection, the 1996 used a fixed sample size. However, there were other significant differences. In the 1987 survey sampling was done from a list of all admissions during the year. Later, to avoid multiple chances of selection some sampled persons were dropped as ineligible, since they had also had a chance of selection as a current resident. Research found this method caused more variation in sample size than the method used in 1996. In 1996, the data collector cleaned the sampling lists to remove persons who had a previous chance of selection in the same facility. In both cases some sample was dropped as ineligible because persons could have been selected from a previous list in another facility. (Bethel et. al., 1993). This was done to simply the calculation of the probability of selection which is almost impossible to calculate if one allows selection in any facility aside from the earliest stay within the year.

Analytic Comparisons

In order to compare the two designs we calculated standard errors and other precision related results for each of the three samples selected for the two surveys. These were: (1) the facility sample, (2) the current resident sample and (3) the admission sample. We did this because each sample has very unique aspects and the comparative differences between results from the two surveys could vary. For instance, using frame reported beds as the measure of size may work well for estimates of current residents, but the quality of results obtained with the same measure of size for the new admission sample could be quite different. For example, after correction for differences in sample sizes, relative standard errors obtained in the 1987 survey for characteristics of admissions are of less quality than those for current residents. (Bethel, et. al., 1993)

For each sample for the two surveys, we calculated results for several sets of variables in 3 or 4 categories of related variables. These categories were:

Facility Sample
- numbers of facility by types, for instance, ownership, region and size
- total beds by facility types, for instance, type of ownership and size
- total residents at the beginning of the year by facility types.

Resident Sample
- resident totals by demographics, for instance, martial status and age
- resident totals by health characteristics, for instance and ability to walk
- resident totals by mental ability, for instance and ability to recognize faces.

Admission Sample
- admission totals by demographics
- admission totals by health characteristics
- admission totals by mental ability
- admission totals during the year by facility types.

There were 61 estimates made with the facility sample, 61 with the resident sample and 80 with the admission sample. The reason there were more admission estimates is that for admissions this sample is required to calculate total admissions estimates. For residents

this estimate is made with the facility sample. This difference is a result of timing of data collection and availability of information within a nursing home. For instance, for each facility in the sample, one collects the number of current residents in the facility as of the beginning of the year. Thus, there was a current resident value for each facility. Thus estimates of the number of residents, say by regions is a facility estimate. For admissions, because persons can be declared ineligible because they had an admission in another facility earlier in the year, the facility value of the admissions in that facility is an overestimate. Thus to make estimates of persons who were admissions during the year, one must determine the sample of persons who were eligible and then sum the person level weights to estimate totals for this group. Thus, for breaks such as, region, the totals for current residents is a facility level estimate, the estimates for persons whose first stay in a facility in the year began during the year is a person level estimate. Hence, we have slightly more person level estimates for the admissions sample. This can be seen from the examples in the table descriptions. Note, the 3rd group under facility is the same type of estimates as the 4th group under admissions.

For each estimate for each survey, using the SUDAAN software for analysis of complex survey data, we calculated the estimates, their sample sizes, standard errors, relative standard errors and design effects. (Shah, et. al., 1993 and Shah, et, al., 1995) We then calculated ratios of adjusted standard errors, adjusted relative standard errors, square roots of design effects and unit standard deviation. The latter two ratios were adjusted for sample size difference by multiplying by the square roots of the ratio of the sample size for each estimate. We developed the latter two comparisons based upon the following approximation:

$$Std\ error\ =\ DE^{.5}\frac{N\sigma}{\sqrt{n}}$$

where
DE is the design effect for the sample
N is the population size
n is the sample size and
σ is the standard deviation for a simple random sample for the population, weighted using the measure of size. (Shah, et. al., 1993)

If one examines this formula, the size of DE reflects the quality of the stratification and sort used in the selection process. The size σ, which we refer to as the unit standard deviation, reflects general variability caused by the population and weights. This includes the quality of the measure of size and the effects of the second phase sampling process on weights. Together, the two measures account for the total differences in the two standard errors after adjusting for sample size.

We were able to calculate values for DE and σ because The SUDAAN software gives an estimate of the design effect. Once, given this estimate and the variance estimate, we used values of N and n available to find σ. The results of this process for the three samples and each type of variable is given below. The values given are averages of the ratios of 1987 values over 1996 values calculated for the two sets of samples and selected variable sets.

Analysis and Comments

For facilities and residents the 1996 design showed the expected improvements in both design effects and unit standard deviations. The results also followed the patterns one might expect. For instance, it appears there was more improvement in design effect for resident variables than estimates for facilities. We would expect that since a key reason for the selection of some strata, such as, hospital based, was because of the big differences in costs and types of patients for this type of facility. Hence, the larger improvement in variables related to resident characteristics.

We also see a pattern in the unit standard deviation comparisons that we would expect. The value for estimates of the set, facilities, is the lowest.

Comparative Quality Indicators									
SAMPLE	**SET**	**RSE**		**SE**		**DE[5]**		**USD**	
		Mean	std err	Mean	std err	Mean	std err	Mean	std err
Facility	Type	1.69	.87	1.20	.30	1.12	.28	1.13	.20
Facility	Beds	1.77	1.23	1.29	.44	1.09	.34	1.22	.26
Facility	Residents	1.77	1.21	1.36	.49	1.12	.36	1.27	.28
Residents	Demographic	1.51	.28	1.37	.27	1.20	.26	1.25	.13
Residents	Mental	2.18	1.50	1.42	.21	1.17	.17	1.32	.16
Residents	Health	1.46	.68	1.38	.24	1.14	.09	1.29	.19
Admissions	Demographic	1.08	.96	0.38	.20	0.53	.23	0.52	.20
Admissions	Mental	0.73	.21	0.36	.14	0.39	.20	0.58	.09
Admissions	Health	0.75	.18	0.34	.10	0.41	.12	0.53	.12
Admissions	Facility Type	1.02	..44	0.44	.22	0.56	.23	0.57	.20

It also is only slightly higher than the projected 10% improvement we felt we would obtain from the change in the allocation method for the second phase of sampling. This should be the key improvement in the estimates of numbers of facilities. The other variables are estimates which are correlated with the number of facility beds. Thus one would expect a further improvement in estimates of numbers of residents and beds for the various breakdowns of the populations. This would occur due to the use of a measure of size that is better correlate with counts of residents. This is seen in the higher improvement in unit standard deviations for the remaining categories of estimates made with the facility and resident samples.

For admissions the 1996 design did not produce the desired results. Work with data from 1987 predicted that the number of beds correlated well with the number of residents and admissions. Unfortunately, in the 1996 survey this did not happen. This was caused by a relatively small yet significant number of facilities in the sample. These were a new type of hospital based facility which did not exist in 1987. This set of facilities contained a limited number of beds, but specialized in short stays. Thus, although they represented approximately 1% of the beds in nation, they represented close to 20% of the admissions. Because of their small per cent of the total beds, but large number of admissions, the strata which contained hospital based facilities had far more variability in admissions data than other strata. However, our allocation proportional to the number of beds essentially assumed otherwise. Thus, our design effects and unit standard deviations for the admissions sample were much worse than for the 1987 survey when these facilities did not exist. We needed to put a disproportionate part of the sample units in the hospital based strata for the variables we studied.

We should note that many of the approximate confidence intervals about the means would include the value of 1. This would indicate that the results are not significant. However, to indicate the strength of these results, for facility and current resident estimates, not once was a design effect or unit standard deviation better in the 1987 data. For admissions, among the 80 comparisons for only one case were the design effects and unit standard deviation better in the 1996 survey.

This, we feel, shows, as a whole, the absolute improvement in the 1996 survey for the facility and current resident estimates and the lack of improvement for the admission estimates.

Presently, data concerning length of stay in nursing homes and costs are not available for estimation. However, we do know that because of the number of admissions per bed in these facilities, that the stays for persons in these types of facilities are very short. Thus, although the number of total admissions in these facilities is disproportionately high the total expenditures in these facilities is likely not to be of the same magnitude. Thus, for these key variables their effects on variances may be of a smaller magnitude.

Having experienced these results in the 1996 survey compared with the 1987 survey for estimates of admissions, we must ask ourselves if this really means that the 1996 sample design is "worse" than the 1987 design. The answer is possibly no. If the 1987 design had been applied to the current world, we may have experienced the same problem or worse.

We have learned several important lessons:

- Past data does not always predict future results, one must carefully consider changes in one's universe. This is something perhaps we tend to overlook because most worlds do not change much.
- The same sample design can be very good for one set of variables and very bad for another.
- There are cases where the old stand by, proportional allocation, is very far from optimal.
- One cannot control for all factors in an evaluation of two sample designs.

References

Bethel, J., P. Ward and G. Kalton (1993). *Alternative Sample Designs for the NMES IPC: Comparing the Cost Effectiveness of Two and Three Stage Sample Designs for the National Medical Expenditure Survey.* Westat, Inc.: Rockville, MD.

Bethel, J. B. and P. Broene (1997). *Final Sample Design Report for the 1996 Medical Expenditure Panel Survey Nursing Home Component.* Westat, Inc.: Rockville, MD.

Cochran, W. G. (1977). *Sampling Techniques.* John Wiley and Sons: New York, NY.

Cohen , S. B. , D. E. B. Potter and P. Flyer (1993). *Sample Design of the Institutional Population Component* (AHCPR Pub. No 94-0017). National Medical Expenditure Survey Methods 6, Agency for Health Care Policy and Research. Rockville, MD: Public Health Service.

Scirroco, A. (1989). Nursing Home Characteristics: 1986 Inventory of Long Term Care Places. *Vital Health Statistics 14(33).* National Center for Health Statistics: Hyattsville, MD.

Scirroco, A. (1994). Nursing Homes and Board and Care Homes. *Advance Data, no. 244.*
National Center for Health Statistics: Hyattsville, MD.

Shah, B. V., R. E. Folsom, L. M. LaVange, K. E. Boyle, S. C. Wheeless and R. L. Williams (1993). *Statistical Methods and Mathematical Algorithms used in SUDAAN.* Research Triangle Institute: Research Triangle, NC.

Shah, B. V., B. G. Barnwell and G. S. Bieler (1995). *SUDAAN: User's Manual.* Research Triangle Institute: Research Triangle, NC.