RECONCILING RESPONDENT REPORTS AND MEDICARE CLAIMS FOR NATIONAL ESTIMATES OF HOSPITAL USE

Sally C. Stearns, Kevin Hayes, Gary G. Koch, University of North Carolina at Chapel Hill
Mary Grace Kovar, Centers for Disease Control and Prevention
Sally C. Stearns, Health Policy and Administration, Chapel Hill, NC

KEY WORDS: Medicare claims, survey reports

Merged data from surveys and administrative files are a valuable analytic resource. One challenge in using these data sets comes in understanding the limitations and possible gaps in the data. Although in theory two sources should yield identical information, the combination of data from multiple sources may result in contradictory reports. While such inconsistencies may not lessen the value of the data, careful review of the merged file is required so that inferences and estimates are reasonable.

An example of such an inconsistency arose from merging Medicare claims data with survey data that included queries about hospital use over the prior year. The Longitudinal Study of Aging (LSOA) is a survey based on a nationally representative sample of 7,527 individuals who were living in the community and who were 70 or older in 1984 (Kovar et al., 1992). The baseline survey was the Supplement on Aging to the 1984 National Health Interview Survey (Fitti and Kovar, 1987). Reinterviews (with the sample member or a proxy) were conducted in 1986, 1988, and 1990. Data from the National Death Index and Medicare claims files were linked with the survey records.

The details of linking the survey data to the Medicare claims files are in Kovar et al. (1992). In brief, 80 percent of the records of LSOA participants were matched to the Master Enrollment File, 8 percent of the sample had no number reported, and 12 percent did not match to the Master Enrollment File.

Comparison of Reported Use and Claims. Hospitalizations reported by the respondent and Medicare claims were analyzed for the 80 percent of the sample that was matched to the Medicare Master Enrollment File. In 1986, 1988, and 1990, sample members (or proxies) were asked for the number of times they had been hospitalized in the year prior to the interview. These data were compared to the number of claims for the interview month and the twelve months prior to the interview, as only the interview month was known. Since the thirteen month period for identifying claims exceeded the twelve- month recall period of the interview, "over-reporting" (reports exceed claims) should be reduced, and "under-reporting" (claims exceed reports) should be increased. Therefore, the difference between "over" and "under" reporting should be a minimum estimate of the true difference.

In contrast to other studies where survey responses were compared with medical records (Marquis, 1984), reported hospitalizations were higher than the claims records showed. Table 1 summarizes the comparisons using dichotomous indicators of whether the sample person had (a) one or more hospital claims and (b) one or more reported hospitalizations for each year. Two points are of note:

- o In each year, reported use exceeded use measured by the claims.
- o The proportion of cases with reported admissions but no Medicare claims increased over time. In contrast, the proportion of cases with no reported admissions but one or more claims stayed roughly constant.

Comparisons in Table 1 show that the rate of inconsistent comparisons was less for self-reports than for proxy reports, and the degree of "over" or "under" reporting was greater in both directions for the proxy reports. However, the rate of "over" reporting was always greater than the rate of "under" reporting, and the difference between "over" and "under" reporting increased each year.

The reported use undoubtedly contains some error. For example, retrospective reports of hospital use may overstate the number of hospitalizations due to telescoping. Telescoping may be particularly likely in the LSOA since the recall period for some of the questions referred to the last time the person was interviewed, an interval of more than one year. However, the consistency with which the claims were less than

the reported use and an understanding of ways in which claims may be missing from Medicare statistical files (HMO enrollment, processing delays, etc.) resulted in a concern that using the claims data to measure use might systematically understate actual utilization. Therefore, we developed procedures to combine and refine the estimates of hospital use.

Reconciling Inconsistencies. The consistently higher recall reports of hospital use provide evidence of a possible problem of missing claims data. Although the difference in the likelihood of hospitalization in 1990 of 4.4 percent between reported use and claims may seem minor, the implications of this difference can be substantial.

This problem can be illustrated using national estimates of the use of hospital services during the year prior to the 1990 interview by 3,473 sample members who: (a) participated in the 1984 LSOA; (b) were considered to be enrolled in Medicare; and (c) had complete interviews in 1990. While this group represents only Medicare enrollees age 76 or older and is therefore only a subset of total hospital users in 1990, the estimates show the magnitude of the error that may occur in projections of hospital use.

As noted earlier, only 80 percent of the sample was matched to the Medicare claims file. A logistic regression revealed that the persons who matched were younger, were more likely to be male, or to be a self-respondent in 1984 than those who did not. In addition, sample persons who had not died by 1990 were more likely to have a completed 1990 interview if they were younger, had attended college, or did not live alone in 1984.

The survey weights were rescaled using a two stage process that first adjusted the weights according to the factors associated with matching to the Medicare files and then rescaled the weights using the factors associated with having a completed interview in 1990. Using the rescaled weights, the 3,473 sample members represent an estimated 12.4 million Medicare enrollees age 76 or older in 1990. This estimate is within five percent of the estimate of 11.9 million Part A enrollees aged 76 or older in 1990 based on data provided by the Bureau of Data Management and Strategy (BDMS) in 1992.

The estimates in Table 2 show that 24.1 percent of people age 76 or older in 1990 had one or more hospitalizations based on reported use and 19.5 percent had one or more hospitalizations based on the claims file. If the reported use is correct, then the difference in these estimates of 4.6 percentage points represents a minimum of 570,834 people among the 12.4 million persons estimated to be alive who had hospital stays that are missing from claims files. Furthermore, these people represent 19.1 percent of persons with reported stays.

Table 2 also shows these estimates according to the type of respondent to the 1990 interview. Again, the reported use measures are higher for each comparison, and the discrepancy between reported use and use measured by claims is smallest for persons who responded for themselves.

Alternative Estimates. Given the inconsistencies between claims and reported use, it may be useful to combine the information to obtain an alternative estimate. For example, if one believes that the self-reported data are the most reliable information on hospital use, then it would be appropriate to develop a correction factor, δ, which is equal to the ratio of odds for self-reported use and claims data. (The correction factor δ is an index for applying the correction, and it should not be confused with the odds ratio.) Specifically, let:

p_s = probability of hospitalization based on self-report for self-respondent

p_c = probability of hospitalization based on claims for self-respondent

Then:

$$(1)\ \delta=\frac{(p_s)/(1-p_s)}{(p_c)/(1-p_c)}$$

The greater the self-reported use is relative to use measured by claims for the self-respondents, then the greater is δ.

The correction factor can then be applied to the odds for the probability of hospitalization for proxy respondents (household or other contact person) to get an "adjusted odds" (AO). Let:

p_p = probability of hospitalization based on claims for person with proxy respondent

Then:

$$(2)\ AO=\delta\times\frac{(p_p)}{(1-p_p)}$$

In effect, the adjusted odds has been increased based on the extent to which the self-report exceeds the claim for the self-respondent

interviews. The adjusted odds can be converted to an adjusted estimate of the probability of hospital use for persons with proxy respondents based on the following formula:

$$(3)\ \textit{Probability of hospital use} = \frac{AO}{AO+1}$$

This formula represents a reverse transformation that results in a proportion that falls within the range from zero to one.

As an example, using the data in Table 2, the correction factor based on the information for self-respondents is:

$$(4)\ \delta=\frac{(19.6/80.4)}{(16.5/83.5)}=1.234$$

Therefore, the estimated probability of hospitalization for sample participants who had a proxy respondent will be based on claims adjusted upwards by approximately 23.4 percent. Table 2 shows the effect of this adjustment on the probability of hospital use for the two types of proxy respondents. The net effect on the estimated probability of hospitalization for the total population is that 23.0 percent of Medicare beneficiaries aged 76 in 1990 or older had one or more hospitalizations during the prior year.

In summary, the essence of the correction is to: (1) assume that self-respondent reports of hospitalizations are true; (2) calculate the ratio of the odds of a self-respondent reporting stays but not having claims in the file for the stays; and (3) estimate the proportion of proxy respondents who have hospital stays by inflating the claims for those with proxy respondents by the ratio.

The calculations are based on the assumption that self-reports provide a "true" measure of hospital use. However, it is possible that the self-reports of hospital use are exaggerated due to telescoping. In adjusting the estimates of hospital use, therefore, it may be preferable to use a weighted average of the self-reports and claims data to calculate a weighted correction factor δ^*, based on a weight ϕ and defined as:

$$(5)\ \delta^* = \frac{\dfrac{(\phi\times(p_s)+(1-\phi)\times(p_c))}{(\phi\times(1-p_s)+(1-\phi)\times(1-p_c))}}{(p_c)/(1-p_c)}$$

The weights ϕ could represent the "data quality" of the various measures, though in practice selection of the weight might be ad hoc. For example, if we believe that the self-reported use is correct two-thirds of the time and that the claims are correct one-third of the time, then ϕ=2/3, δ^*=1.15, and the estimated probability of hospitalization during the prior year for the total population of Medicare beneficiaries aged 76 in 1990 would be 21.8 percent.

Standard Errors and Confidence Intervals. Standard errors and confidence intervals can be calculated for the alternative estimates. Such calculations must address two considerations: the complex survey design, and the fact that the estimate is a composite of the estimates from two different sources.

The first issue was addressed by using SUDAAN, a statistical package that calculates estimates and their standard errors according to survey weights and data structure (Shah, 1991). The estimates for the total column in Table 2 were calculated using SUDAAN with survey weights that were adjusted to reflect differences between sample members included in and excluded from the estimation.

The second issue was addressed by combining information on the standard errors for the two estimates and using the idea of the design effect to obtain an estimate of the standard error of the composite estimate.

The design effect (DE) is defined as the ratio of the variance of a statistic from a complex sample to the variance of the same statistic from a simple random sample of the same size:

$$(6)\ DE=\frac{Variance_{Complex sample}}{Variance_{Simple Random Sample}}=\frac{se^2}{p(1-p)/n}$$

where **p** represents the probability for a binomial variable such as the probability of hospitalization and **se** represents the standard error, and **n** represents the sample size (Landis et al., 1982). Using the information from Table 2, the design effect for the claims estimate is 1.046, while the design effect for the reported-use estimate is 1.071. The larger of the two (the maximum design effect, or MDE) is used to produce a conservative (high) estimate of the standard error for the composite estimate. An estimate of the standard error of the composite estimate can be obtained using the following formula:

$$(7)\ se_{composite} = \sqrt{MDE} \times \sqrt{p_{composite}(1-p_{composite})/n}$$

Composite standard errors calculated using this formula are provided in Table 2 for the alternative estimates. These standard errors are used in Table 3 to obtain confidence intervals for the percent of the Medicare-enrolled population aged 76 or older expected to have one or more hospitalizations in 1990 based on the four estimates of use.

In order to evaluate which measure provides the best estimate, we compared the estimates in Table 3 to an estimate of the number of Medicare enrollees aged 76 or older who had hospitalizations in 1990 based on data provided by the Bureau of Data Management and Strategy (BDMS). Our estimate, based on those data, is that there were approximately 2.96 million enrollees with hospitalizations. (The number of enrollees with hospitalizations was estimated by assuming that four-fifths of the enrollees age 75-79 were age 76-79 using an average of 1989 and 1990 data because the Bureau of Data Management and Strategy could only provide calendar year data for 5-year age intervals. While the estimate from all the interview reports of use comes closest to BDMS's number, the estimate based on the assumption that data from self-respondents are correct may be the best estimate since it includes the BDMS estimate within the 95 percent confidence interval and it is slightly smaller than that number. (The LSOA estimate should be a bit lower than the BDMS number since the LSOA excluded persons who were institutionalized in 1984).

Discussion and Recommendations. While claims data merged with health survey data provide a rich data source, the comparisons presented in this paper have identified some possible gaps in the completeness of such claims data and have highlighted the need for caution in relying exclusively on such data, particularly for making national estimates of hospital use. Key points are:

- o Use of claims data from merged surveys may result in underestimates of use;
- o Retrospective survey reports of use in the LSOA provide an estimate of hospital use that appears to be much closer to actual use than do claims data;
- o One method for obtaining reasonable estimates from merged surveys such as the LSOA may be to increase estimates based on claims by a correction factor based on self-reported use.

The advantages of the correction factor procedure presented here are: it is fairly simple; it does not rely on proxy reports (which do not agree with claims data as well as self-reports do); and it involves essentially negligible bias if having missing claims is a random event. However, if persons with proxy respondents differ from self-respondents in the likelihood of having missing claims, then the correction might be biased. Despite this disadvantage, the method provides estimates that seem reasonable based on comparisons with an estimate of hospitalizations from another source.

The size of the correction factor may be substantial. A correction factor of 23.4 percent based on differences between claims and self-reports appears to be reasonable for adjusting claims-based national estimates from the 1990 LSOA. However, the magnitude of the correction factor will decrease as the completeness of the claims data increases. Since the discrepancy between claims and self-reported use is greatest for 1990 (Table 1), the correction for an earlier year would be less.

While it is not possible to make a recommendation about a correction factor for other datasets, analysts using data from administrative records merged with survey data should be aware that the administrative records may not be more accurate, especially for the most recent time periods. They should investigate their own data and might want to develop their own correction factor. The appropriate factor could change over time and will probably differ among different administrative records. It will certainly depend on such things as the length of time required before the administrative records are complete. For example, Gaumer and Stavins (1992) had to drop three states from their analysis of Medicare claims because the bills had been returned to the intermediary and had not yet been reentered into the central system.

The appropriateness of the correction factor will also depend on whether records are available for all survey participants. For example, there are no claims in the Medicare files for many persons enrolled in HMOs. We did not have information

from the LSOA on who was enrolled in an HMO, although this information could be obtained through a survey. It would also be possible to determine from the survey the number of stays that were not covered by Medicare.

This analysis has focused on implications of missing claims data for making national estimates. In many cases, however, data sets such as the LSOA are used for inferential analyses of relationships between patient characteristics and use of hospital services. While random measurement error in the dependent variable of hospital use does not lead to bias in estimates of the coefficients of explanatory variables, systematic under-reporting of hospital use could cause bias in coefficient estimates. For example, people enrolled in HMOs may, on average, be healthier. Possible corrections for that kind of problem are more complicated than the adjustments discussed in this paper and are not addressed here.

References

Gaumer, GL, J Stavins. (1992) Medicare use in the last ninety days of life. Health Services Research; 26(6):725-742.

Fitti, JE, MG Kovar. (1987) The Supplement on Aging to the 1984 National Health Interview Survey. Vital Health Stat; 1(21).

Kovar MG, M Chyba and JE Fitti. (1992) The Longitudinal Study of Aging: 1984-1990. Vital Health Stat; 1(28).

Landis JR, et al. (1982) A Statistical Methodology for Analyzing Data from a Complex Survey: The First National Health and Nutrition Examination Survey, Vital Health Stat; 2(92).

Marquis, K. (1984) Record checks for sample surveys. In Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines. National Academy Press. Washington, DC.

Shah, BV, BG Barnwell, BG Hunt, P Nileen and LM LaVange. (1991) SUDAAN User's Manual, Release 5.50. Research Triangle Institute, Research Triangle Park, NC.

Funding for this research was provided by the Joint Health Policy Review conducted by the Milbank Memorial Fund and the King's Fund. Emily Fowler provided able research assistance.

Table 1
Comparison of Annual Probability of Hospitalization: Reported Use and Medicare Claims

Comparison Category	Interview Year 1986	1988	1990
Total Sample	3,417	4,240	3,473
a. Percent with Consistent Comparisons	91.5%	90.7%	88.6%
0 claims, 0 reported stays	72.5%	73.8%	72.4%
1+ claims, 1+ reported stays	19.0%	16.9%	16.2%
b. Percent with Inconsistent Comparisons	8.5%	9.3%	11.4%
0 claims, 1+ reported stays ("Over")	4.9%	6.1%	7.9%
1+ claims, 0 reported stays ("Under")	3.6%	3.2%	3.5%
c. Difference ("Over" minus "Under")	1.3%	2.9%	4.4%
Sample with Self-Report	2,630	3,269	2,611
a. Percent with Consistent Comparisons	93.0%	91.9%	90.1%
b. Percent with Inconsistent Comparisons	7.0%	8.1%	9.9%
0 claims, 1+ self-reported stays ("Over")	3.8%	5.3%	6.5%
1+ claims, 0 self-reported stays ("Under")	3.2%	2.8%	3.4%
c. Difference ("Over" minus "Under")	0.6%	2.5%	3.1%
Sample with Proxy Report	787	971	862
a. Percent with Consistent Comparisons	86.5%	87.0%	84.2%
b. Percent with Inconsistent Comparisons	13.5%	13.0%	15.8%
0 claims, 1+ proxy-reported stays ("Over")	8.6%	8.6%	12.0%
1+ claims, 0 proxy-reported stays ("Under")	4.9%	4.4%	3.8%
c. Difference ("Over" minus "Under")	3.7%	4.2%	8.2%

Source: Longitudinal Study of Aging, Version 4
Sample is defined as LSOA participants who matched to Medicare Master Enrollment File and who responded (self or proxy) to the LSOA Survey in the year indicated. The percentages are for the sample and have not been calculated using survey weights.

Table 2
Self versus Proxy Reports of Predicted Hospital Use: Comparisons of Reported Use and Claims Data for 1990

Measures of Hospital Use	Total	Type of Interview: Self-Report	Proxy (Household)	Proxy (Contact)
Sample and Estimated Population				
LSOA Sample Participants	3,473	2,611	695	167
Medicare Enrollees Represented*	12,409,430	9,238,383	2,537,615	633,432
Estimates of Population Use*				
One or More Part A Claim	19.5%	16.5%	27.3%	32.2%
(Standard Error)	(0.688)			
One or More Reported Hospitalization	24.1%	19.6%	36.8%	38.4%
(Standard Error)	(0.751)			
Reconciled (Adjusted) Estimates of Use				
Assumes Self-Report to be Correct	23.0%	19.6%	31.7%	36.9%
(Composite Standard Error)	(0.739)			
Assumes Truth Between Self-Report and Claims**	21.8%	18.6%	30.2%	35.4%
(Composite Standard Error)	(0.725)			

Table 3
National Estimates and Confidence Intervals for Probability of Hospital Use
by Medicare Enrollees age 76 or older in 1990*

	Expected Value	95% Confidence Interval: Lower Bound	Upper Bound
Use Measured by Claims			
Percent with 1 or More Hospitalizations	19.5%	18.3%	20.7%
Persons with 1 or More Hospitalizations	2,419,839	2,270,926	2,568,752
Use Measured by Interview Reports			
Percent with 1 or More Hospitalizations	24.1%	22.7%	25.5%
Persons with 1 or More Hospitalizations	2,990,673	2,816,941	3,164,405
Use Assuming Self-Report to be Correct			
Percent with 1 or More Hospitalizations	23.0%	21.5%	24.4%
Persons with 1 or More Hospitalizations	2,848,144	2,668,504	3,027,781
Use Assuming Truth is Between Self-Report and Claims**			
Percent with 1 or More Hospitalizations	21.8%	20.4%	23.2%
Persons with 1 or More Hospitalizations	2,706,603	2,530,192	2,883,015

Notes for Tables 2 and 3

Source: Longitudinal Study of Aging, Version 4

* The sample for the estimated population of 12.4 million enrollees includes persons who participated in the 1984 LSOA, were determined to be enrolled in Medicare, and responded (self or proxy) to the 1990 interview. The survey weights were rescaled for nonmatches to the Master Enrollment File and incomplete 1990 survey responses.

**Reported use is weighted at 2/3, and claims use is weighted at 1/3.