Testing a diagnosis-related group index for skilled nursing facilities

by Philip G. Cotterill

Interest in case-mix measures for use in nursing home payment systems has been stimulated by the Medicare prospective payment system (PPS) for short-term acute-care hospitals. Appropriately matching payment with care needs is important to equitably compensate providers and to encourage them to admit patients who are most in need of nursing home care. The skilled nursing facility (SNF) Medicare benefit covers skilled convalescent or rehabilitative care following a hospital stay. Therefore, it might appear that diagnosis-related groups (DRG's), the basis for patient classification in PPS, could also be used for the Medicare SNF program. In this study, a DRG-based case-mix index (CMI) was developed and tested to determine how well it explains cost differences among SNF's. The results suggest that a DRG-based SNF payment system would be highly problematic. Incentives of this system would appear to discourage placement of patients who require relatively expensive care.

Introduction

In the past several years, case-mix measures for nursing homes have been developed in an attempt to reflect differences in the resource intensity of the care needed by different types of nursing home patients and to incorporate these differences into payment systems. (For a review of the literature on this subject, see Stassen and Bishop, 1983). Interest in prospective payment systems has fostered work on case-mix measures. Prospective rates based on average costs without adjustments for case-mix differences would encourage nursing homes to admit only those patients with the lowest resource needs. Therefore, access would be difficult for the patients most in need of nursing home care, who often require relatively expensive care.

The Medicare prospective payment system (PPS) for short-term acute care hospitals has reinforced interest in case-mix-adjusted payment for nursing homes. PPS creates an incentive for hospitals to discharge Medicare patients sooner than they would have been discharged under retrospective cost-based reimbursement. Reductions in hospital lengths of stay may also result in an increased number of sicker Medicare beneficiaries who need nursing home care.

For several reasons, the patient classification system of diagnosis-related groups (DRG's), which serves as the basis for payments in PPS, is a logical candidate as the case-mix measure for a prospective payment system for the Medicare skilled nursing facility (SNF) program. First, the SNF Medicare benefit covers skilled convalescent or rehabilitative care following a hospital stay of at least 3 days. It thus seems appropriate to classify the SNF stay in the same terms as the hospital stay. Second, a common classification system for hospital and SNF care would greatly facilitate the eventual development of an inclusive payment for an episode involving both types of care. Third, administration of a Medicare SNF prospective payment system would be simpler and less costly if each SNF patient could be assigned to a DRG prior to admission to the SNF. Any case-mix measure that required patient assessment after admission to the SNF would be very expensive. It would involve consistent and accurate collection of potentially detailed primary data. Moreover, Medicare SNF patients are spread across more than 5,000 SNF's, the majority of which have very few Medicare patients.

However, there are reasons to question the feasibility of basing Medicare SNF payment on DRG's. Evidence exists that diagnosis is not a strong predictor of differential use of nursing home resources for Medicare SNF patients (Health Care Financing Administration, 1985). It has been shown that the care needs of nursing home patients arise primarily from the need for assistance in performing basic activities of daily living, such as eating, dressing, bathing, and toileting. Research does indicate that Medicare SNF patients differ from other nursing home patients in ways that are consistent with Medicare's focus on short-term convalescent or rehabilitative care (Shaughnessy et al., 1985). However, it is not known whether functional dependency varies systematically across DRG's. Also, functional dependency may vary as much within as across DRG's. The presence of multiple diagnoses for a patient may also limit the usefulness of any single diagnosis in predicting utilization of SNF resources. In short, the relative importance of case mix and intercorrelations among its various dimensions remain unsettled questions.

These issues were explored by testing the extent to which a DRG-based case-mix index (CMI) explains differences in cost among SNF's. For reasons related to the nature of the data available (discussed in detail later), the CMI's tested in this study are based on a DRG derived from the SNF admission diagnosis rather than the hospital discharge diagnosis. Medicare data for 1980 indicate that these diagnoses differ for two-thirds of SNF stays. As a result, this article does not directly bear on the issue of using the hospital DRG as the basis for SNF payment. Rather, the more general issue is addressed of the applicability of DRG's as a classification system for SNF payment.

The main components of the study were construction of DRG-based CMI's for SNF's and

Reprint requests: Philip G. Cotterill, 2-B-14 Oak Meadows Building, 6325 Security Boulevard, Baltimore, Maryland 21207.

multiple regression analysis of SNF costs, including the SNF CMI's as explanatory variables. In the analysis of SNF costs, two issues were addressed: the degree of explanatory power of the SNF CMI and the magnitude of the effect of the CMI on SNF costs. Both issues have implications for the use of SNF DRG's as the basis for case-mix adjustments in a potential Medicare SNF prospective payment system.

First, if the SNF CMI does not have explanatory power comparable with that of the hospital CMI and of variables that serve as proxy measures of SNF case mix, its usefulness as an SNF payment variable is highly questionable.

Second, and equally important, the magnitude of the cost differences associated with differences in the SNF CMI would determine the incentives for SNF's to admit patients with different levels of care needs. If SNF's were paid a fixed rate for each DRG (subject to adjustments for wage differences), an SNF's payments would be proportional to its CMI. (An SNF would receive 10-percent higher payments if its CMI were 1.10 than if it were 1.00.) If, however, an SNF's costs do not vary proportionally with its CMI, there will be a financial incentive to avoid certain types of patients. In particular, if costs are disproportionately higher for higher CMI's, SNF's will have an incentive to avoid patients with greater than average care needs. For example, suppose that an SNF with a CMI of 1.10 had costs 15 percent higher than they would be if its CMI were 1.00. The additional cost (15 percent) exceeds the additional payment (10 percent), and the SNF has no financial incentive to accept the more intensive case mix.[1] Appropriately matching payment with care needs is important to equitably compensate providers and to encourage them to admit patients who are most in need of SNF care.

The article is organized as follows. First, the chief tasks involved in the construction of the SNF CMI's are described and descriptive statistics are presented, including a comparison with the hospital CMI used in PPS and described by Pettengill and Vertrees (1982). Next, findings of the analysis of SNF costs are presented. A variant of the hospital cost equation that Pettengill and Vertrees (1982) used to test the hospital CMI was estimated, and the SNF results were compared with the hospital findings. With some adaptation, the SNF cost equation estimated by Sulvetta and Holahan (1986) was reestimated, adding SNF CMI measures to the equation. Finally, there is a discussion of the implications of the study results for the use of DRG's as the basis for case-mix adjustments in a potential Medicare SNF prospective payment system.

SNF case-mix index construction

The construction of the SNF CMI involved three tasks.

- A DRG based on the SNF admission diagnosis was assigned to every Medicare SNF admission in calendar year 1980 that involved at least 1 covered day.
- Relative weights were constructed for SNF DRG's to measure the expected resource intensity of cases in each DRG relative to the average case.
- CMI's were computed for each SNF by multiplying each DRG's relative weight by the proportion of an SNF's total cases in each DRG, then summing these products over all DRG's.

DRG assignment

The "DRG grouper" software used in PPS was modified so that a DRG could be assigned to each 1980 Medicare SNF stay. Three data elements were used in determining SNF DRG's: the diagnosis recorded as the reason for admission to the SNF, an indicator of whether additional diagnoses were recorded, and the patient's age.

The SNF DRG may not be the same as the DRG associated with the hospital stay that qualified the patient for admission to the SNF for the following reasons:

- The SNF admission diagnosis may differ from the hospital principal diagnosis.
- The SNF stay record contains no information about surgical procedures performed during the hospital stay, an important determinant of the hospital DRG.
- The SNF and hospital stay records may differ as to whether additional diagnoses were present or recorded.
- The patient's age may differ on the SNF and hospital stay records.

All of these differences may occur either for good cause or because of inaccuracies in the data. Because of these differences, the issue of combining hospital and SNF payments for episodes of care is not directly addressed in this article. Instead, the more general issue of DRG's as a case-mix measure for SNF's is explored.

It is possible to determine the hospital DRG for the roughly 20 percent of SNF stays that match the hospital stays contained in the Medicare provider analysis and review (MEDPAR) data files. The chief reason for not using those files is that few SNF's have enough MEDPAR cases to create a reliable CMI. As a result, too few observations would be available for the cost analysis.

DRG's were assigned to approximately 220,000 SNF stays in 5,157 SNF's. The number of stays ranged from more than 42,000 in DRG 236 (fractures of the hip and pelvis) to about 500 cases in DRG 202 (cirrhosis and alcoholic hepatitis), the least frequent DRG used in the study.

[1] An algebraic exposition of the relationships among payments, costs, and the CMI is available from the author on request.

DRG relative weights

Relative weights measure the expected resource consumption of each SNF DRG relative to the resource consumption of the average case across all DRG's. Two simplifications were adopted in deriving the SNF DRG relative weights. First, primarily for computational ease, the analysis was restricted to the 66 most frequent DRG's, which comprise over 90 percent of the SNF stays. Second, charges rather than costs were used as the proxy measure of resource consumption.

Charges were used for several reasons. They are available on the SNF stay file, whereas costs are more difficult to compute. Missing data on ancillary cost-to-charge ratios would have necessitated the elimination of many SNF's and significantly reduced the number of cases that could be used to calculate relative weights. Finally, charge weights have been shown to correspond very closely to cost weights (Cotterill, Bobula, and Connerton, 1986) and are now used in the Medicare PPS (*Federal Register*, 1985).

Separate weights were constructed on a per-day and a per-case basis. The day is the traditional unit of payment for nursing home care, and individual variations in length of stay may make case payment impractical. However, there is little if any research in which the feasibility of per-case payment for nursing homes is tested. Additionally, there is interest in eventually linking SNF payment with the per-case payment of hospitals under PPS. Therefore, it was elected to construct per-case relative weights and CMI's.

Six sets of relative weights were constructed based on accommodations, ancillary, and total charges per day and per case. Separate weights were computed for accommodations and ancillary charges to determine whether SNF DRG's affect these components of care differently. The accommodations weights probably underestimate cost variations among SNF DRG's because relatively few SNF's set daily rates based on the nursing and other routine services required by individual patients. Routine costs, which correspond to accommodations charges, represent about 80 percent of total costs, so the total charge weights are also expected to underestimate the variation in total cost among SNF's. Only covered charges and covered days were used in the calculations.

All charge variables were standardized for geographical wage rate differences. Some cases were eliminated on the basis of extreme values for accommodations or ancillary charges per day. After inspecting the means and standard deviations of these variables for each DRG, cases were excluded if wage-adjusted accommodations charges per day were less than $20 or more than $200 or wage-adjusted ancillary charges per day were more than $200. Cases with covered days in excess of the maximum 100 days were also excluded. These exclusions were made to eliminate unreasonable data values that, in all likelihood, result from errors in coding the data. These criteria resulted in the elimination of 3,301 cases, or 1.6 percent of all cases in the top 66 DRG's.

Relative weights were computed by dividing each average charge variable for each DRG by the corresponding average charge for all cases or days for all 66 DRG's. Examination of the resulting relative weights reveals that:

- The range of values among DRG's is smaller for SNF's than it is for hospitals. The minimum SNF weight is .348, and the maximum is 1.51. (Both are weights for ancillary charges per case.) Weights for total charges per case range from .519 to 1.423, compared with a range of .346 to 7.48 among weights for hospital total operating costs. Weights for total charges per day range only from .791 to 1.283.
- Per-case weights display greater dispersion than per-day weights because of variation in length of stay as well as in charges per day.
- Ancillary weights display greater dispersion than accommodations weights. The standard deviation of the per-day accommodations weights is only about 8 percent of the mean, compared with a standard deviation of 20-30 percent for the ancillary weights.

SNF case-mix indexes

Six CMI's were computed for each SNF represented in the 1980 Medicare SNF stay file. CMI's were calculated by multiplying the appropriate relative weight for each SNF DRG by the proportion of the SNF's cases or days in each DRG, then summing these products for the 66 DRG's. Of the 5,118 SNF's, 28 percent (1,442) had more than 50 stays, 48 percent (2,472) had 11-50 stays, and 24 percent (1,204) had 10 stays or less. However, because virtually all 1980 cases in these DRG's were included, the CMI's are not subject to sampling error. The CMI's are representative of SNF cases in the 66 DRG's even when the total number of cases is small.

The number of DRG's represented in any SNF's CMI ranges from 1 to 64. About one-third of the SNF's had cases in more than 20 DRG's, and another one-third had cases in 10-20 DRG's. The value of an SNF's CMI is not highly correlated with the number of DRG's represented. The Pearson correlation coefficients range from $-.07$ for the CMI based on accommodations charges per case to $+.05$ for the CMI based on ancillary charges per day.

How well CMI's explain cost variation among SNF's depends more on the extent of variation among CMI's than on the variation among relative weights. However, it is useful to note that the CMI's are consistent with the relative weights. For example, most SNF's diversify their caseload among several DRG's; thus, the CMI's are expected to vary less than the SNF DRG relative weights do. Indeed, the standard deviations of the CMI's are only one-half to

one-third as large as those of the relative weights. Further, the pattern of dispersion among the six different CMI's is similar to that among the six sets of relative weights.

- The per-case CMI's display greater dispersion than do the per-day CMI's. The standard deviations of the per-case CMI's, which range from .07 to .13, are from 1½ to 2 times as large as those of the per-day CMI's.
- The CMI's based on ancillary charges vary more than those based on routine and total charges. The standard deviations of the CMI's for ancillary charges range from .09 to .13 and are roughly two to three times larger than those for routine and total charges.

SNF CMI's are less dispersed than hospital CMI's are. The standard deviation of the CMI based on total charges per case is .08 for SNF's, compared with .10 for hospitals. The smaller variation in SNF CMI's is plausible because hospitals span a much greater range of complexity in technology and procedures than do SNF's.

The greater the variation in a CMI, the greater its potential for explaining cost differences among SNF's. In Table 1, the means and standard deviations of the six CMI's are presented for all 5,118 SNF's. Information on the 3,410 SNF's that could be identified from Medicare cost report data is displayed by selected characteristics known to be associated with significant cost differences: hospital-based or freestanding status and urban or rural location. Hardly any variation exists in the average CMI's of SNF's with these characteristics. Because the costs of hospital-based SNF's are approximately twice those of freestanding SNF's, it is interesting that their CMI's differ very little. In contrast, 1981 Medicare hospital data show considerable variation by cost-related characteristics. The average CMI for rural hospitals was .96, compared with 1.04 for urban hospitals and 1.14 for hospitals with large teaching programs.

In their study of SNF costs, Sulvetta and Holahan (1986) found the percent of an SNF's days provided to Medicare patients to be an important predictor of cost. The CMI's for SNF's are shown in Table 2 according to the percent of Medicare days: less than 10 percent, 10-35 percent, 36-65 percent, and more than 65 percent. The variation in the CMI's across the four Medicare groups is small. However, the pattern is consistent for all CMI's: a higher mean CMI is associated with a higher percent of Medicare days. The range in means across the Medicare groups varies from 1 percentage point for CMI 1 (accommodations charges per day) to 5 percentage points for CMI 2 (ancillary charges per day) and 7 percentage points for CMI 5 (ancillary charges per case).

Despite the limited variation in the SNF CMI's (Tables 1 and 2), the results are consistent with previous research. The findings of Shaughnessy et al. (1985) and Sulvetta and Holahan (1986) suggest that a higher proportion of Medicare days is positively related to the intensity of case mix. Lack of variation in the CMI's between hospital-based and freestanding SNF's is also consistent with the Shaughnessy et al. (1985) finding that most of the case-mix differences between hospital-based and freestanding SNF's were attributable to the presence of sicker non-Medicare patients in hospital-based SNF's. These facilities differ relatively little in the types of Medicare patients served.

Analysis of SNF costs

As stated earlier, two previous cost function analyses (Pettengill and Vertrees, 1982; Sulvetta and Holahan, 1986) served as the point of departure for this investigation of the relationship between the SNF CMI's and SNF costs. In this section, the basic results of the regression analysis of SNF costs are presented. Next the implications of these results are explored to answer the following questions:

- Is the explanatory power of the SNF CMI strong enough to be a useful payment variable?
- How would the size of the SNF CMI impact on costs affect placement of heavy- and light-care SNF patients in a potential Medicare SNF payment system based on SNF DRG's?

Hospital cost function

The first test of the explanatory power of the CMI's involved estimating the SNF analog of the equation for hospital cost per case used by Pettengill and Vertrees (1982) to test the hospital CMI. In the hospital equation, cost is assumed to depend on bed size, the Bureau of Labor Statistics hospital wage index, location (rural and three city-size variables), and the CMI. The hospital cost equation serves as a useful comparison because, for the most part, it contains the variables used in PPS to adjust payments.

Instead of estimating one equation for total operating costs per case, six equations and six dependent variables were estimated corresponding to each of the SNF CMI's: routine, ancillary, and total costs both per day and per case. The only variable from the hospital equation not in the SNF equation is the extent of the hospital's teaching activity. As in Pettengill and Vertrees (1982), all continuous variables are expressed in logarithms.

Each of the six equations was estimated five times: for all SNF's and for each of the four Medicare groups shown in Table 2. Separate equations were used for the Medicare groups because the SNF CMI was expected to have greater explanatory power for SNF's with a higher percent of Medicare days. The SNF CMI is based only on Medicare cases, whereas the cost variables pertain to all patients cared for in Medicare-certified beds. Therefore, the higher the percent of Medicare days, the better the correspondence between the CMI's and the cost variables. Moreover, the separate equations provide additional information about the relationship between the CMI's and the percent of Medicare days variable, which may itself be a proxy measure of case mix.

Table 1

Mean and standard deviation of case-mix indexes (CMI's) for Medicare-certified skilled nursing facilities (SNF's), by type of SNF and CMI: United States, 1980

CMI	All SNF's	Type of SNF			
		Freestanding		Hospital based	
		Urban	Rural	Urban	Rural
Number of SNF's[1]	5,118	2,320	621	234	235
CMI 1: Accommodations charges per day	.999	.999	.997	.998	1.00
	(.027)	(.026)	(.022)	(.023)	(.029)
CMI 2: Ancillary charges per day	.983	.987	.981	.985	.974
	(.093)	(.091)	(.084)	(.091)	(.101)
CMI 3: Total charges per day	.995	.996	.993	.995	.994
	(.037)	(.036)	(.031)	(.033)	(.039)
CMI 4: Accommodations charges per case	.997	.998	.996	.999	.995
	(.071)	(.070)	(.068)	(.067)	(.078)
CMI 5: Ancillary charges per case	.987	.991	.986	.992	.974
	(.135)	(.132)	(.130)	(.138)	(.135)
CMI 6: Total charges per case	.995	.997	.993	.997	.990
	(.084)	(.082)	(.081)	(.082)	(.088)

[1] *The 4 types of SNF's total to 3,410 rather than 5,118. Information on type of SNF comes from the Medicare cost report file, which includes only 3,410 of* the 5,118 SNF's in the stay file. The number of SNF's shown here also differs from the number of SNF's reported in Tables 2-6. Tables 2-6 reflect the number of SNF's included in regressions, and the elimination of missing and statistical outlier data accounts for the smaller number used.

Table 2

Mean and standard deviation of case-mix indexes (CMI's) for Medicare-certified skilled nursing facilities (SNF's), by percent of Medicare days and CMI: United States, 1980

CMI	Total	Percent of Medicare days			
		Less than 10	10-35	36-65	More than 65
Number of SNF's	2,817	1,887	602	192	136
CMI 1	.999	.997	.999	1.004	1.009
	(.023)	(.025)	(.016)	(.015)	(.019)
CMI 2	.987	.976	1.004	1.014	1.029
	(.080)	(.088)	(.054)	(.049)	(.045)
CMI 3	.996	.992	1.000	1.006	1.014
	(.031)	(.034)	(.020)	(.019)	(.019)
CMI 4	.997	.989	1.014	1.014	1.017
	(.062)	(.065)	(.050)	(.052)	(.053)
CMI 5	.991	.973	1.022	1.026	1.042
	(.118)	(.126)	(.091)	(.093)	(.085)
CMI 6	.996	.985	1.016	1.017	1.023
	(.073)	(.077)	(.059)	(.061)	(.059)

NOTE: See Table 1 for definitions of CMI's and explanatory footnote.

The overall explanatory power of these equations (R^2), is presented along with the regression coefficients of the CMI's and their t statistics in Table 3. These results do not compare favorably with the results for hospital costs. Pettengill and Vertrees (1982) explained 72 percent of the variation in hospital's total operating costs per case. In contrast, as seen in Table 3, the highest R^2 for any SNF equation is .28 (total cost per day for SNF's with more than 65 percent Medicare days). R^2 values of less than .10 are common.

However, the results for the individual regression coefficients of the SNF CMI's are considerably better than the overall explanatory power of the equation. Most coefficients are statistically significant at the .05 level or better. The largest t statistic shown in Table 3 is 10.6, which is highly significant. Nevertheless, the results are weaker than the hospital results, for which the CMI coefficient had a t statistic greater than 25.

The size of the CMI's effect on costs tends to increase as the SNF's percent of Medicare days increases. The size of these coefficients is particularly striking. For example, the coefficient for total cost per day of SNF's with more than 65 percent Medicare days is 11.45. This implies that a 10-percent difference in the CMI would be associated with a 114.5-percent difference in total cost per day.

Table 3

R^2 values, case-mix index (CMI) regression coefficents, and t statistics for hospital cost equation[1] adapted for skilled nursing facilities (SNF's), by percent of Medicare days: United States, 1980

Equation	Total	Percent of Medicare days			
		Less than 10	10-35	36-65	More than 65
Number of SNF's	2,817	1,887	602	192	136
Routine cost per day and CMI 1:					
R^2	.01	.01	.06	.04	.13
Regression coefficient	*1.80	.63	*3.09	*4.82	*5.41
t statistic	3.31	1.00	2.50	2.06	2.98
Ancillary cost per day and CMI 2:					
R^2	.16	.03	.03	.03	.00
Regression coefficient	*1.90	.20	*1.44	*2.94	*4.68
t statistic	6.52	.70	2.18	1.97	1.96
Total cost per day and CMI 3:					
R^2	.12	.05	.13	.18	.28
Regression coefficient	*1.07	*-.54	*3.42	*5.63	*11.45
t statistic	4.88	2.68	6.25	4.53	6.33
Routine cost per case and CMI 4:					
R^2	.04	.03	.16	.14	.10
Regression coefficient	*1.99	*1.23	*3.23	*3.31	*1.40
t statistic	10.4	5.23	8.92	5.06	2.21
Ancillary cost per case and CMI 5:					
R^2	.09	.02	.00	.00	.00
Regression coefficient	*2.27	*1.16	.89	*2.42	1.60
t statistic	8.10	3.96	1.39	1.96	.78
Total cost per case and CMI 6:					
R^2	.05	.02	.15	.13	.10
Regression coefficient	*1.77	*.974	*2.75	*3.00	*1.50
t statistic	10.6	4.90	8.75	5.31	2.44

*Significant at least at .05 level.

[1](Pettengill and Vertrees, 1982).

NOTE: The regression coefficient indicates the percentage difference in cost associated with a 1-percent difference in the CMI. See Table 1 for definitions of CMI's and explanatory footnote.

Three conclusions can be drawn from the results shown in Table 3.

- The statistical significance of the SNF CMI's indicates that a systematic positive relationship exists between CMI's and SNF costs.
- The variables that are important in explaining cost differences among hospitals explain relatively little of the total cost variation among SNF's.
- In most cases, the size of the impact of the SNF CMI on costs is quite large and indicates that SNF costs vary more than proportionally with differences in the CMI. According to these elasticity estimates, an SNF with a CMI greater than 1.0 would have costs in excess of its DRG payments, whereas an SNF with a CMI lower than 1.0 would receive payments that exceed its costs. The effect of these disparities would be to discourage SNF's from accepting Medicare patients with above average care needs. Setting SNF DRG payment rates could thus present serious problems of perverse placement incentives.

SNF cost function

The second test of the ability of the CMI's to explain variation in SNF costs involved reestimating the equations reported in Sulvetta and Holahan (1986), adding the SNF CMI to their equation. In their equation, cost is assumed to be a function of several factors. Bed size, rural-urban location, wage index, occupancy rate, type of ownership, whether the SNF belongs to an investor-owned chain, whether the SNF is hospital based, and whether the SNF is a distinct-part facility are all included, as are two proxies for case mix: percent of Medicare days and admissions per bed. These variables explain about 55 percent of the total variation in cost. When dichotomous variables for each State are added, the equation explains 64 percent of the variation in total cost per day and 67 percent of the variation in routine cost per day. The most highly significant variables are percent of Medicare days, hospital-based status, occupancy rate, type of ownership, and wage index. In contrast to the hospital equation, this equation contains several structural characteristics that contribute significantly to the explanation of cost variation but are not likely candidates for incorporation into a payment system.

In reestimating the Sulvetta and Holahan (1986) equations, as with the hospital cost equation, routine, ancillary, and total costs both per day and per case were used as dependent variables. Again, separate equations were run for each of the four Medicare groups. To facilitate comparison of these results with those of Sulvetta and Holahan (1986), the equations were not estimated in logarithmic form.

Table 4

R^2 values, case-mix index (CMI) regression coefficients, and t statistics for skilled nursing facility (SNF) equation[1] adapted for use with CMI's, by percent of Medicare days: United States, 1980

Equation	Total	Percent of Medicare days			
		Less than 10	10-35	36-65	More than 65
Number of SNF's	2,817	1,687	602	192	136
Routine cost per day and CMI 1:					
R^2	.54	.36	.52	.61	.72
Regression coefficient	*.33	-.02	*.88	*2.15	*2.71
t statistic	2.80	-.13	2.99	2.72	2.42
Ancillary cost per day and CMI 2:					
R^2	.50	.08	.19	.31	.31
Regression coefficient	*.06	.00	.07	*.71	*1.41
t statistic	3.03	.29	1.57	3.96	3.42
Total cost per day and CMI 3:					
R^2	.60	.33	.52	.58	.67
Regression coefficient	*.18	*.24	*1.46	*2.85	*4.18
t statistic	1.50	2.32	5.26	3.20	2.29
Routine cost per case and CMI 4:					
R^2	.24	.22	.33	.40	.39
Regression coefficient	*23.91	*12.60	*45.14	*43.81	*18.54
t statistic	8.29	3.73	6.87	3.76	1.30
Ancillary cost per case and CMI 5:					
R^2	.34	.05	.12	.16	.10
Regression coefficient	*2.40	*.82	*2.66	*17.52	10.02
t statistic	4.89	1.97	2.40	4.84	1.24
Total cost per case and CMI 6:					
R^2	.28	.20	.29	.33	.28
Regression coefficient	*23.62	*11.02	*46.13	*66.10	61.59
t statistic	7.94	3.37	6.63	4.59	.49

*Significant at least at .05 level.

[1](Sulvetta and Holahan, 1986).

NOTE: The regression coefficient indicates the dollar difference in cost associated with a 1-percent difference in the CMI. See Table 1 for definitions of CMI's and explanatory footnote.

The results of reestimation of this equation are presented in Table 4. The R^2 values for the logarithmic equations in Table 3 are not directly comparable with those for the linear equations in Table 4 because the total variation in cost differs when the cost variable is transformed into logarithms (Rao and Miller, 1971). However, the resultant differences in R^2 should not be large enough to confound the basic comparisons of interest.

The overall explanatory power of the equations shown in Table 4 is much greater than that of the equations shown in Table 3. R^2 values for the per-day regressions in Table 4 range from a high of .72 for routine cost per day for SNF's with more than 65 percent Medicare days to a low of .08 for ancillary costs for SNF's with less than 10 percent Medicare days. R^2 values of equations for total cost per day range from .33 for SNF's with less than 10 percent Medicare days to .67 for SNF's with more than 65 percent Medicare. The explanatory power of the per-day regressions is typically about twice that of the per-case regressions.

The results shown in Table 4 for the individual CMI coefficients are consistent with those shown in Table 3. Most coefficients are statistically significant at the .05 level. In addition, the size of the CMI coefficient tends to be greater in SNF's with a larger percent of Medicare days.

The CMI coefficients shown in Tables 3 and 4 should be interpreted differently. The continuous variables in the hospital equation are in logarithms, so the CMI coefficients in Table 3 are elasticity estimates; that is, they indicate the percentage difference in cost associated with a 1-percent difference in the CMI. The coefficients shown in Table 4 are simple linear estimates. They indicate the dollar difference in cost associated with a difference in the CMI of 1 percentage point. However, the coefficients in Table 4 can be used to compute elasticities for mean values of cost and the CMI's. These calculations yield elasticities greater than 1, a finding that is consistent with the results shown in Table 3 and indicates that SNF costs vary more than proportionally with the CMI. In general, the results shown in Table 4 reinforce the conclusion drawn from Table 3 regarding the strength and size of the positive relationship between the CMI's and SNF costs.

Explanatory power of SNF CMI

Do the results of the multiple-regression cost analysis imply that the explanatory power of the SNF CMI is great enough to make it a useful variable for SNF payment? The statistical significance of the SNF CMI regression coefficients shown in Tables 3 and 4 generally indicates (with a probability of .95 or more)

that the estimated positive relationship between the CMI and SNF costs is not merely a chance occurrence. However, statistical significance alone is not sufficient to answer the question. No definitive standards exist, so two criteria are applied:

- How does the explanatory power of the SNF CMI compare with that of the hospital CMI?
- How does the explanatory power of the SNF CMI compare with that of the two proxy variables for Medicare SNF case mix, percent of Medicare days and admissions per bed?

Impact on R^2 values

To assess the relative explanatory power of the SNF and hospital CMI's, the CMI was added to the cost equations and the resulting changes in their explanatory power (as measured by changes in the R^2) were compared. It is necessary to compare changes rather than R^2 alone because independent variables other than the CMI also contribute to overall explanatory power.

The change in R^2 that occurs when the relevant SNF CMI is added to the basic equation of Sulvetta and Holahan (1986), excluding State variables, is shown in Table 5. Increases in R^2 attributable to the addition of the CMI range from zero for all SNF's combined and for the group with less than 10 percent Medicare days to 5-10 percentage points for several per-case CMI's for SNF's with more than 10 percent Medicare days. The maximum increase in R^2 for any per-day equation is 6 percentage points for ancillary costs per day in the groups with 36-65 percent and more than 65 percent Medicare. An increase of 2 percentage points is common in the per-day equations.

To test whether the regressions reported in Table 5 might understate the explanatory power of the SNF CMI, the CMI was added to a version of the Sulvetta and Holahan (1986) equation that does not contain the two case-mix proxy variables, percent of Medicare days and admissions per bed. The resulting changes in R^2 were examined. Deleting these two variables reduces the R^2 significantly in a majority of cases but has very little impact on the change in R^2 obtained by adding the CMI to the equation. These effects are very similar to the ones shown in Table 5. As a result, they are not reported in this article.

A similar comparison was made for hospitals: adding the hospital CMI to the equation for total operating costs per case and observing the change in the R^2. Adding the CMI to the set of independent variables (bed size, wage index, rural and city-size location variables, and teaching commitment) increases the R^2 for the equation from .69 to .72.

This difference of 3 percentage points is comparable with the changes shown in Table 5 for SNF's with a high percent of Medicare days. Thus, in this comparison, the SNF CMI appears to have as much explanatory power as does the hospital CMI. However, there is an important difference between the effects of the hospital CMI and the SNF CMI. The R^2 increase of 3 percentage points attributed to the hospital CMI understates its explanatory power. The hospital CMI and the wage index alone explain approximately 68 percent of the variation in hospital operating costs per case. However, the SNF CMI and the wage index explain at most 28 percent of SNF total costs per case (for SNF's with more than 65 percent Medicare days). Of course, this result indicates that the hospital CMI is correlated with other variables that are themselves highly correlated with hospital costs.

Correlations with other independent variables

To pursue this point further, hospital and SNF CMI's were regressed on the independent variables in their respective cost equations. For hospitals, the CMI was regressed on bed size, rural and urban variables, wage index, and teaching activity. The R^2 of this equation was .42, and the bed-size variable was particularly important.

For SNF's, each CMI was regressed on the variables in the Sulvetta and Holahan (1986) equation. For all SNF's combined, regardless of percent of Medicare days, the R^2 values for these equations range from .03 to .05. When the sample is restricted to SNF's with more than 10 percent of Medicare days, some R^2 values improve, but they never exceed .12. Clearly the SNF CMI's are much less highly correlated with other independent variables than is the hospital CMI.

In the hospital case, the high degree of intercorrelation makes it difficult to precisely identify how much cost variation is attributable to case mix. However, large urban teaching hospitals treat a more complex mix of cases than is found in small, rural nonteaching hospitals. Therefore, the high intercorrelation provides some assurance that the CMI is measuring what it is intended to measure.

In the SNF case, the intercorrelations are much lower and are only slightly suggestive of relationships consistent with other evidence. In certain subgroups of SNF's, some statistically significant positive relationships occur among various SNF CMI measures and the variables hospital-based status and percent of Medicare days. (The correlation between the CMI and percent of Medicare days was apparent in Table 2, but the correlation between the CMI and hospital-based status was not apparent in Table 1.) In addition, a statistically significant negative relationship exists between some CMI measures and proprietary ownership. However, in none of these cases is the relationship comparable with the degree of intercorrelation observed among the hospital variables.

Unfortunately, there are two ways of interpreting the SNF results. On the one hand, the effect of the SNF CMI might represent the true case-mix effect, which is not highly correlated with other SNF characteristics. On the other hand, the SNF CMI's low degree of intercorrelation with other variables provides little assurance that it is accurately measuring the true case-mix effect. Given the lack of other corroborative evidence, it is not clear what effects the SNF CMI is capturing.

Table 5

R^2 values for skilled nursing facility equation with and without addition of case-mix indexes (CMI's), by percent of Medicare days: United States, 1980

Equation	Total	Percent of Medicare days			
		Less than 10	10-35	36-65	More than 65
Routine cost per day:					
R^2 without CMI 1	.54	.36	.52	.59	.71
R^2 with CMI 1	.54	.36	.52	.61	.72
Ancillary cost per day:					
R^2 without CMI 2	.50	.08	.19	.25	.25
R^2 with CMI 2	.50	.08	.19	.31	.31
Total cost per day:					
R^2 without CMI 3	.59	.33	.49	.56	.65
R^2 with CMI 3	.59	.33	.52	.58	.67
Routine cost per case:					
R^2 without CMI 4	.22	.21	.28	.36	.38
R^2 with CMI 4	.24	.22	.33	.40	.39
Ancillary cost per case:					
R^2 without CMI 5	.34	.05	.11	.06	.09
R^2 with CMI 5	.34	.05	.12	.16	.10
Total cost per case:					
R^2 without CMI 6	.26	.20	.24	.26	.28
R^2 with CMI 6	.28	.20	.29	.33	.28

NOTE: All R^2 statistics are adjusted for degrees of freedom. See Table 1 for definitions of CMI's and explanatory footnote.

Further tests of explanatory power

To further assess the extent to which cost differences are associated with CMI differences, the CMI regression coefficient was multiplied by the standard deviation of the CMI. The result represents the cost difference associated with a difference in the CMI of one standard deviation. Dividing this amount by the standard deviation of the relevant cost variable produces the beta coefficient, a relative measure of the variation in cost associated with variation in the CMI.

These computations were performed for each of the six CMI's and their corresponding cost variables. The CMI regression coefficients were based on the linear SNF equation used for Table 4, which was reestimated pooling the data for all SNF's with more than 10 percent Medicare days. Also, the standard deviations of CMI and cost variables for all SNF's with more than 10 percent Medicare days were used in computing the beta coefficients. The resulting cost differences associated with the CMI range from $1.93 (ancillary cost per day) to $5.08 (total cost per day) and from $66.31 (ancillary cost per case) to $277.64 (total cost per case.) The relative measures range from 10 percent (routine cost per day) to 13 percent (total cost per day) and from 14 percent (ancillary cost per case) to 21 percent (routine cost per case). An analogous computation for hospital total operating cost per case yields a relative CMI effect of 22 percent. Clearly, the cost variation associated with the SNF CMI's tends to be less than that associated with the hospital CMI used in PPS.

Finally, the explanatory power of the SNF CMI was compared with that of the two proxy case-mix variables, percent Medicare days and admissions per bed. As in the previous case, the beta coefficients were based on the regression coefficients and standard deviations of all variables for SNF's with more than 10 percent Medicare days. Even within this group, percent of Medicare days accounts for about twice as much variation in cost as does the SNF CMI. (The beta coefficients are roughly double those of the SNF CMI variables.) However, considerably less cost variation is typically accounted for by admissions per bed than by the per-day SNF CMI's. The per-case comparisons can be ignored because there is, by definition, a fairly high negative correlation between costs per case and admissions per bed. (Admissions per bed and length of stay are approximate reciprocals, and length of stay is positively correlated with cost per case.)

Discussion

In summary, this examination indicates that the SNF CMI explains somewhat less cost variation than does the hospital CMI and considerably less than does percent of Medicare days, a proxy case-mix variable. The differences between SNF and hospital CMI's are interesting and difficult to interpret. The high correlation between the hospital CMI and hospital size provides some indication that the CMI is measuring what it is intended to measure. The lack of any comparably strong correlation between the SNF CMI and any other identifiable SNF characteristic offers no

such comfort. Two plausible possibilities exist:

- SNF case-mix differences may be largely uncorrelated with the other factors associated with cost differences. The variety among State Medicaid programs and the different orientations of various types of nursing home ownership may have created a patchwork quilt of care patterns that bears little relation to Medicare SNF case mix.
- Limitations of the data and of the applicability of DRG's may reduce the explanatory power of the SNF CMI measures and bias upward the SNF CMI elasticity estimates.

Data limitations are most likely to involve errors in measurement of the SNF CMI and omitted variable effects. For example, measurement errors may compress the observed values of the CMI. Compression means that for above average CMI values, the observed CMI associated with a given cost value may be too low; for below average CMI values, the observed CMI may be too high. Also, some omitted independent variables (functional dependence, idiosyncratic patterns of care, and/or non-Medicare SNF case mix) may be positively correlated with the SNF CMI. If so, the cost values observed at above average levels of the CMI would be higher than they should be, and cost values observed at below average levels of the CMI would be lower than they should be.

Compression of the CMI's is likely to result from the fact that the charge data used in constructing the DRG relative weights are unlikely to capture differences in resource use among patients within the same facility who are classified in different DRG's. Compression may result from any source of variation in case mix that affects accommodations charges (routine costs). Failure of these charges to reflect patient differences in functional dependence may be especially important. Patient-level costing and pricing are not common, and when they are attempted, Medicare patients typically fall into a single category at the high end of the scale. Compression is likely to be particularly serious because routine costs account for roughly 80 percent of total costs.

Further tests were conducted to determine possible omitted variable effects. To test the possible positive correlation between idiosyncratic patterns of care and the SNF CMI, the SNF CMI was correlated with the number of nursing hours per day. The Pearson correlation coefficients range from .06 (routine cost per case) to .11 (total cost per day). These statistics suggest that this relationship could account for at least part of the upward bias in the SNF CMI elasticity estimates. When the staffing variable is added to the cost equation, the SNF CMI coefficient declines slightly in some cases, and in others it increases. No strong conclusions could be drawn from this test.

Throughout the regression analysis, the effects of non-Medicare patients on the cost variables were controlled for by partitioning the data into four Medicare groups and by including percent of Medicare days as a continuous variable. Non-Medicare patients affect the cost variable used in the analysis because the variable is a measure of the average cost of all patients treated in Medicare-certified beds. Medicare patients tend to be more expensive to care for than other patients are, so the observed cost variable may understate the cost for Medicare patients by an amount that increases as the percent of Medicare patients in the SNF decreases.

Finally, the SNF cost equations were rerun separately for hospital-based and freestanding SNF's to determine whether hospital-based SNF's might account for the disproportionate cost effect of the SNF CMI. The dichotomous hospital-based variable included in the combined cost equation might fail to completely account for the effect of a more costly non-Medicare case mix in hospital-based SNF's. This possibility was suggested by the finding of Shaughnessy et al. (1985) that much of the case-mix difference between hospital-based and freestanding SNF's could be attributed to the more intensive care needs of non-Medicare patients in hospital-based SNF's. The results did not support the hypothesis. The largest CMI coefficient is for freestanding SNF's in the group with more than 65 percent of Medicare days. It is almost as large as is the coefficient in the equation for hospital-based and freestanding SNF's combined.

These problems of omitted variables and measurement error can never be fully addressed by data of the type used in this study. Problems such as compression of the DRG relative weights and the failure to incorporate a measure of functional dependency can be addressed only by research involving primary data collection.

Conclusion

In this article, the development is described of a case-mix index for SNF's based on DRG's defined primarily by the diagnosis recorded as the reason for admission to the SNF. Tests of the relationship between SNF costs and this SNF CMI support the conclusion that DRG-based SNF payment would be highly problematic. A statistically significant positive relationship exists between the SNF CMI and SNF costs. However, the explanatory power of the relationship is generally weaker than the explanatory power of the relationship between the Medicare hospital CMI and hospital costs. Further, the relationship between SNF costs and the SNF CMI is weaker than the relationship between SNF costs and percent of Medicare days, a proxy SNF case-mix measure. These results strongly suggest that the relationship between the SNF CMI and SNF costs is not strong enough to make the SNF CMI a useful payment variable.

Finally, the estimated SNF CMI regression coefficients indicate that SNF costs vary much more than proportionally with the SNF CMI. These large positive elasticities imply that a payment system based on the relative weights developed in this study would have adverse effects on SNF admission practices. Such

a payment system could discourage the admission of Medicare patients with greater than average care needs.

Acknowledgments

The author gratefully acknowledges the data processing assistance of Ronald Prihoda and Janet O'Leary of the Bureau of Data Management and Strategy.

References

Cotterill, P., Bobula, J., and Connerton, R.: Comparison of alternative relative weights for diagnosis-related groups. *Health Care Financing Review.* Vol. 7, No. 3. HCFA Pub. No. 03222. Office of Research and Demonstrations, Health Care Financing Administration. Washington. U.S. Government Printing Office, Spring 1986.

Federal Register: Medicare program: Changes to the inpatient hospital prospective payment system and fiscal year 1986 rates. 50(170), Sept. 3, 1985.

Health Care Financing Administration, Office of Legislation and Policy: *Report to Congress: Study of the Skilled Nursing Facility Benefit Under Medicare.* Jan. 1985.

Pettengill, J., and Vertrees, J.: Reliability and validity in hospital case-mix measurement. *Health Care Financing Review.* Vol. 4, No. 2. HCFA Pub. No. 03149. Office of Research and Demonstrations, Health Care Financing Administration. U.S. Government Printing Office, Dec. 1982.

Rao, P., and Miller, R. L.: *Applied Econometrics.* Belmont, Calif. Wadsworth Publishing Company, 1971.

Shaughnessy, P. W., Kramer, A. W., Schlenker, R. E., and Polesovsky, M. B.: Nursing home case-mix differences between Medicare and non-Medicare and between hospital-based and freestanding patients. *Inquiry* 22(2):162-177, Summer 1985.

Stassen, M., and Bishop, C. E.: Incorporating Case Mix in Prospective Reimbursement for SNF's Under Medicare: Critical Review of Relevant Research. Unpublished paper. Brandeis University Center for Health Policy Analysis and Research, Mar. 1983.

Sulvetta, M. B., and Holahan, J.: Cost and case-mix differences between hospital-based and freestanding nursing homes. *Health Care Financing Review.* Vol. 7, No. 3. HCFA Pub. No. 03222. Office of Research and Demonstrations, Health Care Financing Administration. Washington. U.S. Government Printing Office, Spring 1986.