Case mix, quality, and cost relationships in Colorado nursing homes

by Robert E. Schlenker and Peter W. Shaughnessy

The analyses reported in this article assessed the cost, case mix, and quality interrelationships among Colorado nursing homes. A unique set of patient-level data was collected specifically to measure case mix and quality. Case mix was found to be strongly associated with cost, accounting for up to 45 percent of the variation in cost per patient day. The relationship between quality and cost was weaker; quality variables accounted for only about 10 percent of the cost per day variation. Case mix was also associated with several facility characteristics found to be significant in other cost studies, suggesting that such facility characteristics serve as partial proxy measures for case mix.

The cost-case mix relationships appear to be strong enough to justify incorporating case mix directly in nursing home reimbursement systems. In contrast, the weaker cost-quality association implies that it may not (yet) be appropriate to incorporate quality directly in reimbursement.

Introduction

Objectives

The analyses had two major objectives. The first objective was to assess the degree that nursing home cost is associated with nursing home case mix and quality of care. Such an assessment is important for long-term care reimbursement policy formulation, since a strong association between, case mix or quality and cost would suggest that these variables should be more directly incorporated in Medicare and Medicaid nursing home reimbursement policies.

The second objective was to determine if case mix and quality were associated with certain facility characteristics that have been found to relate to cost in other studies. The existence and magnitude of such associations would provide at least a partial explanation for the cost differences among different types of facilities (such as, between forprofit and nonprofit nursing homes). This information is also important to policy formulation, so that policymakers can decide on the degree that differential payments are warranted on the basis of such facility characteristics (Stassen and Bishop, 1983).

The cost function approach

Nursing home costs are influenced by many factors, and in broad terms, a nursing home cost function can be specified as:

C = f(CM, Q, FC, MKT),

where,

C = cost (usually per patient day),
CM = case mix,
Q = quality,
FC = facility characteristics,

and

MKT = market factors.

The CM and Q variables reflect differences in the "product" produced by the nursing home; the FC variables take into account other facility-specific reasons for differences in cost per patient day among nursing homes (such as size, to measure possible economics of scale); and the MKT variables include factors exogenous to the facility that might influence its costs (such as strong competition resulting from the ready availability of substitute services).

Once individual variables and a functional form are specified, statistical techniques can be used to empirically estimate the parameters of the cost function. The approach expressed in this article was somewhat different from the usual method of placing heavy emphasis on the coefficient estimates for individual variables. Our greatest interest was in the explanatory power (measured by the R^2) of the various categories of variables, particularly case mix and quality variables.

Earlier studies have estimated nursing home cost functions, and most have used available secondary data. Unfortunately, secondary data usually provide only limited measures of case mix and quality, such as whether or not a nursing home is classified as a skilled (SNF) or intermediate care (ICF) facility, and the percentages of patients classified in each category (as case-mix measures), or the deficiencies reported in Medicaid certification and licensure surveys (as a quality measure). The limited availability of case mix and quality data has also prevented most cost studies from separating the effects of facility characteristics, such as ownership or urban or rural location, from the effects of unmeasured case mix and quality variables. For example, nursing home costs are usually negatively associated with the proportion of Medicaid patients. Although that finding may result from greater efficiency being forced on nursing homes by low Medicaid payment rates, it may also be because of nursing homes lowering case mix and quality. Similarly, nonprofit nursing homes typically have higher costs than forprofit homes, and this may also reflect differences in either efficiency or case mix and quality.

This research was supported by the Department of Health and Human Services, Grant No. 18-P-97145/8.

Reprint requests: Dr. Robert E. Schlenker, University of Colorado Health Sciences Center, Center for Health Services Research, Container C-241, 4200 East Ninth Avenue, Denver, Colorado 80262.

In an earlier article, Bishop summarized the results of several cost function studies, and highlighted the lack of adequate case mix and quality measures (Bishop, 1980). A more recent article by Birnbaum (Birnbaum, Bishop, Lee, and Jensen, 1981) reports on several studies using a nationwide nursing home sample, as well as, samples from Massachusetts, New York, and Indiana. Where possible, these studies included "product characteristics" such as certified level of care and services provided, as well as, patient descripters of age, diagnosis, level of dependence in activities of daily living (ADL's), and mental impairment. Detailed quality measures were unavailable. An updated review noting the same difficulties was recently published by Palmer and Cotterill (Palmer and Cotterill, 1983). Additional recent discussions of the strengths and weaknesses of the cost function approach can be found in Bishop (1983) and Scanlon and Weisert (1983). This article attempts to overcome the limitations of secondary case mix and quality measures by using patient-level data and constructing facility-level case mix and quality variables.

The Colorado setting

Comparisons with national data from the 1977 national nursing home survey indicated that Colorado nursing homes were typical in terms of occupancy rate (an average rate of 92 percent), ownership (73 percent for profit), and urban and rural location (35 percent in standard metropolitan statistical areas (SMSA's). However, Colorado nursing homes also differed from the national averages in these respects:

- The percent of Medicaid patients was higher in Colorado (70 percent versus 60 percent).
- The percent of Medicare patients was lower (.4 percent versus 2.3 percent).
- Colorado facilities were slightly larger than the national averages (only 13 percent of Colorado nursing homes had fewer than 50 beds, compared to 42 percent nationally).

The Colorado Medicaid nursing home reimbursement system in 1980 was similar to many State systems. A separate per-diem payment rate was prospectively determined for each nursing home, which was based on the facility's reported costs for a prior 6-month period (adjusted for inflation). A ceiling payment rate was established each year, and was set at the 90th percentile of the total cost per day for all Medicaid patients. If a nursing home's cost per day was below the ceiling, it received an "efficiency incentive," add-on, in its subsequent prospective rate. The add-on was equal to half the amount by which the nursing home's actual cost per day was less than the ceiling. (A maximum was placed on the efficiency incentive—equal to 7 percent of the ceiling.) The Colorado reimbursement system incorporated both features to encourage efficiency and to at least partially provide higher payments to higher cost facilities. Under the assumption that case mix and quality affect costs, the costs associated with an intense case mix or higher quality would be at least partially covered by a higher Medicaid payment rate; this would presumably reduce the reluctance of some facilities to accept such patients or to provide higher quality care. This meant that variation in case mix and quality across facilities was more likely than under other types of payment systems (such as a flat rate approach), and such variation is necessary for cost function analysis. Similarly, the efficiency features built into the system suggested that facility costs should be near the minimum possible for any given case mix and quality combination. (The possibility that minimizing cost was a less powerful objective for nonprofit nursing homes was also tested in the analysis.)

Taking all these factors into consideration, it is our opinion that although Colorado's 1980 nursing homes and reimbursement system were not identical with all the other States, Colorado results can be generalized to other States.

Samples, variables, and methods

Samples

Although the analyses presented used 1980 data, the overall study analyzed the 1978-80 period. Therefore, the facility samples were selected based on 1978 data, and focused on Medicaid-certified, general nursing care facilities. In 1978, 198 nursing homes were in operation in Colorado, of which 167 were certified by Medicaid, and classified as general nursing care facilities. Of these, 10 had to be dropped because they experienced significant organizational changes (for example, closure) or were financially unstable between 1978 and 1980. The remaining 157 facilities were termed the secondary sample, because information was available on these facilities only from secondary sources.

For the patient-level case mix and quality data, a subsample of facilities was used because of the high cost of primary data collection. A total of 74 nursing homes was selected as the primary sample. Although considerable effort was made to ensure randomness for the primary sample, this was not totally achieved because the final selection had to be based on overall data availability and the nursing homes' agreement to participate.

The nursing homes in the primary sample were compared to the remaining facilities in the secondary sample by using certain selected characteristics. These characteristics are: occupancy rate, percent of patients who were Medicaid recipients, number of beds, type of control (forprofit or nonprofit), SMSA designation, certification level (SNF, ICF or a combination of both), and type of facility (freestanding or hospital-based). Table 1 presents these comparisons. None of the differences was signficant at the .05 level. Control type was closest to significance, $p = .06$. The proportion of forprofit facilities was lower in the primary sample (66 percent versus 80 percent). This was actually an advantage for the analyses in that it provided a greater number of nonprofit observations. The chance of the higher proportion of nonprofit facilities biasing the results was dealt with by including a forprofit or nonprofit variable in some of the

Table 1

Comparisons of selected nursing home characteristics for the primary sample and the remaining secondary sample facilities: 1980

Nursing home characteristics	Primary sample		Secondary sample		Significance [1]
	Number	Percent	Number	Percent	
Occupancy rate	74	92.7	83	90.9	[2] .152
Percent Medicaid patients	74	69.5	83	70.9	[2] .305
Facility bed size					
Total	74	100.0	83	100.0	.102
0-60 beds	28	37.8	25	30.1	
61-120 beds	26	35.1	43	51.8	
121 beds or more	20	27.0	15	18.1	
Type of control					
Total	74	100.0	83	100.0	.060
Forprofit	49	66.2	66	79.5	
Nonprofit	25	33.8	17	20.5	
SMSA designation					
Total	74	100.0	83	100.0	.303
Rural	29	39.2	26	31.3	
Metropolitan	45	60.8	57	68.7	
Certification level					
Total	74	100.0	83	100.0	.863
Skilled	23	31.1	29	34.9	
Intermediate	14	18.9	14	16.9	
Skilled and intermediate	37	50.0	40	48.2	
Type of facility					
Total	74	100.0	83	100.0	.983
Freestanding	65	87.8	73	88.0	
Hospital-based	9	12.2	10	12.0	

[1]Exact p-value for the chi-square test (unless otherwise noted).
[2]Exact p-value listed for the Wilcoxon two-sample test is presented since the variable was not normally distributed (as indicated by the Kolmogorov-Smirnov test).

SOURCES: Colorado Department of Social Services, Denver, Colorado: Medicaid Cost Reports, 1980.
Colorado Department of Health, Denver, Colorado: Directory of Licensed Health Facilities, 1980.

regressions and by separately examining the association between control type and the case mix and quality variables.

Hospital-based facilities had more medically intense case mix than freestanding nursing homes (Schlenker, Shaughnessy, Yslas, 1983). Hospital-based facilities' costs were also significantly higher, but this was partly because of cost accounting procedures required of such facilities under Medicaid in Colorado (which parallel Medicare rules). Hence, the inclusion of hospital-based facilities could overstate the relationship between case mix and cost pertinent to only freestanding facilities. For this reason, separate analyses were conducted for the total primary sample of 74 facilities and for the subsample of the 65 freestanding facilities.

Random samples of patients were selected in each primary sample facility for case-mix data collection, and quality data were obtained for a subsample of these patients (a subsample was used because of the lengthy data collection process required for the quality data). Private pay as well as publicly financed (mainly Medicaid) patients were included. Case-mix data were obtained for 1,750 patients and quality data for 912 of those patients.

Variables

The four categories of cost, case mix, quality, and facility characteristics and market variables were included in the analyses. The cost measures were the dependent variables, and the independent variables of case mix and quality were viewed as reflecting "product differences" among nursing homes. Facility characteristics and market factors were independent variables hypothesized to reflect differences in nursing home objectives and efficiency, and in the supply and demand environment faced by each home.

Cost variables

Cost data were obtained from Colorado Medicaid cost reports and measured as cost per patient day. The analyses focused on Medicaid "adjusted" costs, which reflected the outcome of Medicaid auditing procedures intended to remove nonallowable and nonpatient care related expenses.[1] In the Colorado

[1]On the average, the auditing procedures reduced the costs reported by nursing homes by approximately 4 percent. However, the correlation coefficient between adjusted and unadjusted total cost per patient day was .95 for the 1980 data, suggesting that the relationships between cost and factors such as case mix and quality would be essentially the same whether adjusted or unadjusted costs were utilized.

Medicaid cost reports, the major categories are administrative and general, room and board, health care, and property costs. Within the health care category, most expenditures represent nursing staff—nurses and aides—wages and salaries. On the average, these nursing costs comprise more than 90 percent of the health-care cost category and approximately 40 percent of the total cost category.

The two cost variables analyzed were the total- and nursing-cost per patient day. Total cost was selected because all components of nursing home costs were potentially affected by case mix and quality. Nursing cost was hypothesized to be the category most directly affected; hence, case mix and quality were expected to have greater explanatory power in the nursing cost than in the total cost equations.

Some nursing home cost studies have included nonproperty or operating-cost per patient day (i.e., total minus property costs), under the rationale that property costs such as depreciation, interest, and lease payments, are not directly related to patient care. Although analyses using operating-cost per day were carried out, the results were similar to those presented for total cost; therefore, they are not presented here.

Case-mix variables

Various sets of case-mix variables were used in the analyses, including measures of diagnoses, long-term care problems, and ADL's. The case-mix data were collected by the Center for Health Services Research nurses experienced in long-term care and trained especially for this task. Each primary sample facility was visited, and patient-level data were collected from available records (medical charts) and interviews of the primary care providers (nursing staff). Patient observations were made on a limited basis for validation purposes.

Medical diagnoses for each patient were measured by 18 major diagnostic groupings (such as neoplasms, circulatory system disorders, and digestive system disorders). Because medical diagnoses are typically of limited value in long-term care studies (because long-term care patients often have a large number of diagnoses, and services are usually targeted at specific types of problems rather than at specific diagnoses), the long-term care problem measures were developed and include bowel and urine incontinence, immobility, visual impairment, hypertension, urinary tract infection, depression or apathy, and mental retardation or developmental disability (Polesovsky and Shaughnessy, 1982). The measures were developed for 27 long-term care problems, and drew on the work of other researchers and the assistance of an expert panel. The problems were categorized into the four major groups of nursing, communicative, medical, and psychosocial problems, according to the major type of care required. For each patient in the sample, data were collected on the presence and severity of each of the 27 problems.

The ADL indicators are standard case-mix measures in long-term care. They measure whether a patient is independent or dependent in certain functional areas as bathing, dressing, eating, and toileting. Data on 7 ADL measures were obtained for each patient, and several degrees of dependence were measured for each ADL.

The patient-level case-mix data were aggregated across the sample patients in each facility. Each diagnosis, problem, and ADL variable was dichotomized at the patient level, and then aggregated to the facility level to reflect the percentage of the facility's patient sample with each diagnosis, problem, or ADL dependency.

At the patient level, severity scales were also analyzed for the problem variables, as well as, different numbers of dependency categories for each ADL. However, for the facility analyses, the dichotomous variants were found to perform as well as, the more complex variables.

Because many of the case-mix variables measured similar conditions, their intercorrelations were fairly high. Several different groups of case-mix measures yielded similar results when their association with cost was analyzed. The results presented use the following four case-mix variables, reflecting the percent of each facility's patient sample with urinary tract infection, mental retardation and developmental disability, dependence in toileting, and digestive system disorder.

The first two variables are long-term care problem variables; the third is an ADL; and the fourth is a diagnosis. It is important to note that these are illustrative variables, used to demonstrate the overall explanatory power of case mix in the cost functions Because of the intercorrelations mentioned earlier, these variables should be considered as indicator or "tracer" case-mix variables. For any operational reimbursement system, case-mix measures that better represent a patient's total condition would be needed. However, the number of such measures should be fairly small to be practical, and the interactive effects among individual case-mix variables would have to be considered.

Quality variables

The quality measures developed were based on the appropriateness of services provided for each of the 27 long-term care problems previously described. The long-term care problems included or subsumed most relevant functional areas (ADL's), and diagnostic categories important in long-term care. The results of other research and the aforementioned long-term care expert panel were utilized to develop recommended services, service frequencies, and providers for each of the 27 problems. For a subset of patients in each facility and a random sample of five problems for each patient (less than five if the patient had fewer problems), data were collected on specific services provided, frequencies, and providers. Quality scores were then calculated to reflect the extent that each recommended service was provided by the appropriate provider and with the recommended frequency. For each problem, providers and services were weighted on the basis of their relative importance in treating the particular problem as judged by the expert panel.

The resulting scores ranged from 0-100 percent. The quality measurement methodology, as originally developed (Shaughnessy, Breed, and Landes, 1982) was refined for this study (Yslas, Brown, and Shaughnessy, 1982).

To render the quality scores as independent of case mix as possible, each problem-level quality score (i.e., the score for each problem for each patient) was calculated in standard Z-score form. The overall mean was subtracted from each problem score, and the result was divided by the standard deviation of all quality scores for that problem across the entire patient sample. These Z-scores were then aggregated to the patient and facility levels, resulting in quality scores that were relatively independent of case mix. The facility-level aggregate scores were used in the analyses presented, and are referred to as the "quality Z-scores." Extensive reliability tests were conducted on these and the case mix variables (Polesovsky and Shaughnessy, 1982; Ysales, Brown, and Shaughnessy, 1982).

As was done for case mix, analyses were carried out for several different sets of quality variables. The illustrative results presented are typical of those for various sets of quality measures (especially, since the intercorrelations among many of the individual quality variables were fairly high). The indicator quality measures used the Z-scores for two specific long-term care problems—subcutaneous or complicated skin condition, and immobility.

As mentioned earlier in the discussion of case-mix measures, any operational reimbursement system that incorporates quality variables would have to use more comprehensive measures that better reflect overall quality. Further, the quality measures examined are strictly process measures and do not deal with patient outcomes.

Facility characteristic and market factor variables

The facility characteristics that were hypothesized to be associated with differences in nursing home behavior and cost were:

- Control type (forprofit/nonprofit)—that measures potential differences in cost structures as a function of organizational objectives.
- Number of beds and occupancy rate—that test for economies of scale and efficient use of plant capacity.
- Percent of nursing costs representing nursing pool use—that captures the effect on costs of the higher wage rates of nursing pool personnel (pool nurses are brought in to fill temporary staff shortages).
- Fiscal reporting period ending date—that adjusts for possible cost differences due to inflation, because the nursing homes in the sample had different reporting periods.
- Percent of Medicaid patients in a facility—that captures possible differences in behavior resulting from the degree that a nursing home is dependent on Medicaid as a payer.
- Health Care Financing Administration Wage Index (derived from Section 223 routine cost limit tables)—which is a rough indicator of wage rate differences across the State.
- Medicaid certification levels (ICF, SNF, or both)—which are used as proxy case-mix measures and as possible indicators of differences in cost structures resulting from certification requirements (such as nursing hours required for each skilled care patient).

Other, lesser facility characteristics included the separate identification of nonprofit government facilities, the percent of Medicare patients, whether a nursing home was part of a chain of facilities, an indicator of change in ownership during the year, and whether the cost data were based on audited or unaudited cost reports.

Several variables were also included as market factors to test for possible cost impacts of differences in the demand for nursing home care across the State and in the supply of substitute or complementary services. The overall hypothesis was that weak demand and or strong competitive forces would require greater cost containment efforts by facilities. Conversely, strong demand and or weak competitive forces would allow for greater "organizational slack" and higher cost, if all other factors are equal. These variables were measured either at the county or SMSA levels and included:

- Percent of the population aged 65 and over and per capita income as demand factors.
- Nursing home beds per thousand population aged 65 and over as a measure of market supply relative to potential demand.
- Home health visits for the same population group as a potential substitute service.
- The area average hospital occupancy rate as an indicator of potential competition by hospitals for patients.
- The ratios of hospital beds and physicians to the total population as indicators of possible derived demand for nursing home care.

Because many of these market factor variables were associated with the urban or rural categories, variables measuring whether or not a facility was in an SMSA and or in the Denver SMSA were also included as substitute measures of market factor differences.

To reduce the number of facility characteristics and market factor variables (because of the relatively small facility samples with primary case mix and quality data), preliminary regression analyses were conducted on the larger secondary sample of 157 facilities. In these analyses, only proxy case mix and quality variables could be utilized, such as the percentage of skilled care patients as a case mix proxy, and the nursing hours per patient day as a quality proxy. The purpose of these analyses was to identify the major facility characteristics and market factors that were empirically associated with the cost variables. The results indicated that the market factor variables and most of the facility characteristics were of little statistical significance. The three facility

characteristics that were highly significant statistically were forprofit or nonprofit control, the percent nursing pool use, and the percent of Medicaid patients.[2] These variables were included in the analyses for the primary sample of facilities. Further, for the sample of 74 facilities, a variable indicating whether a nursing home was hospital-based or freestanding was also included because of its importance as noted earlier.

Final model and statistical techniques

The final model included the following variables:

Dependent variables

- Nursing cost per patient day.
- Total cost per patient day.

Independent variables

- *Case mix*—the percent of each facility's patient sample with urinary tract infection, mental retardation or developmental disability, dependence in toileting, and digestive system disorder.
- *Quality*—Z-score for each facility's patient sample for the long-term care problems of subcutaneous or complicated skin condition, and immobility.
- *Facility characteristics*—these include: for-profit or nonprofit (0/1, 1 = nonprofit), percent nursing pool use, percent Medicaid patients, and hospital-based or freestanding (0/1, 1 = hospital-based).

A linear functional form was used for the regressions presented. Such a functional form assumes that each variable is independent of the others and that their effects are additive. Other functional forms were explored, such as using interaction terms for combinations of case-mix variables and for the case mix and quality indicators for the same problem, but these approaches added little to the results.

The main statistical techniques employed were correlation and ordinary least-squares, stepwise regression analyses. The individual associations between pairs of dependent and independent variables were examined through correlation analysis, and multivariate associations between the dependent variables and sets of independent variables were assessed using regression analysis. Due to the relatively small sample sizes, only a few independent variables (usually five to ten) were included in any one regression equation. Many combinations of variables were used to determine the various interaction patterns among independent variables, and their individual, as well as, combined effects on the dependent variables.

Because the major objective of the regression approach was to determine the explanatory power of different sets of independent variables (measured by the R^2), and less so to identify statistically significant variables, regression equations were estimated relating the dependent variables to the following independent variables:

- Case mix only.
- Quality only.
- Combinations of case mix and quality.
- Case mix, quality, and facility characteristics combined.

In this way, upper and lower bounds were determined for the explanatory power of the various sets of variables. For example, the maximum explanatory power for case mix was determined as the R^2 when only case-mix variables were included. The minimum explanatory power was the additional R^2 contributed when case mix variables were added to a regression that already included the quality and facility characteristic variables.

Results

Explanatory power of case mix, quality, and facility characteristics

Table 2 presents illustrative regression equations relating the nursing and total cost variables to case mix. For each dependent variable, the results of two regression analyses are presented—one using the 65 freestanding facilities and the other using the 74 freestanding and hospital-based facilities combined. These regressions test the hypothesis that costs vary as function of only case mix.

The results suggest that case mix alone is associated with one-third of the variation in nursing cost per day among freestanding facilities ($R^2 = .33$). As expected, the R^2 is lower for total cost in this sample ($R^2 = .25$). Owing to the greater variation in both case mix and cost when the hospital-based nursing homes are included with the freestanding nursing homes, the R^2's for both dependent variables are higher for the sample of 74 facilities, and for this sample are approximately the same ($R^2 = .43$ and $.45$).

In the nursing cost equations in both samples, the urinary tract infection and digestive system disorder variables had positive and statistically significant coefficients ($p < .05$ using a two-tailed t-test). The toileting variable was highly significant in the sample of 65 facilities but only marginally significant ($p = .083$) in the sample of 74 facilities. However, a one-tailed test hypothesizing a positive association between dependence in toileting and cost would also yield significance in this case. Although the coefficient for the mental retardation and developmental disability variable was negative but insignificant, the negative sign is consistent with the general finding that a higher proportion of such patients was associated with lower cost. In view of the difficulty of measuring quality in this area, it is not clear whether such lower costs reflect lower care needs or lower quality of care, and further research is definitely warranted.

The total cost results are similar to those for nursing cost in each sample. Taken together, these four regressions point to a relatively strong association between patient case mix and nursing home costs.

[2]The certification variables (ICF, SNF, or SNF and ICF combined) tended to be significant when used as proxy case-mix measures in preliminary analyses, but were insignificant when the direct case-mix measures from the primary sample of facilities were also included in the regressions.

Table 2

Illustrative regressions relating nursing and total cost per patient day to case mix: 1980

Variables	65 freestanding nursing homes		74 freestanding and hospital-based nursing homes	
	Coefficient	Significance [1]	Coefficient	Significance [1]
Nursing cost per patient day [2]				
Independent variables, percent of patients with:				
Urinary tract infection	.156	.011	.249	<.001
Mental retardation and/or developmental disability	−.019	.178	−.022	.152
Dependent in toileting	.051	<.001	.041	.083
Digestive system disorders	.054	.016	.052	<.001
Constant	6.710	<.001	6.853	<.001
Total cost per patient day [3]				
Independent variables, percent of patients with:				
Urinary tract infection	.188	.140	.734	<.001
Mental retardation and/or developmental disability	−.056	.059	−.060	.122
Dependent in toileting	.072	.001	.596	.658
Digestive system disorders	.056	.230	.101	.003
Constant	20.909	<.001	20.326	<.001

[1]The level of significance is based on a two-tailed t-test.
[2]For freestanding nursing homes, the mean of the dependent variable is $10.14, and R^2 is .33 ($p < .001$). For freestanding nursing and hospital-based nursing homes, the mean of the dependent variable is $10.47, and R^2 is .43 ($p < .001$).
[3]For freestanding nursing homes, the mean of the dependent variable is $25.26, and R^2 is .25 ($p < .001$). For freestanding nursing and hospital-based nursing homes, the mean of the dependent variable is $27.29, and R^2 is .45 ($p < .001$).

SOURCES: University of Colorado Health Sciences Center, Center for Health Services Research: Primary Study Data, 1980. Colorado Department of Social Services, Denver, Colorado: Medicaid Cost Reports, 1980.

Table 3

Illustrative regressions relating nursing cost per patient day to case mix and quality: 1980

Variables	65 freestanding nursing homes		74 freestanding and hospital-based nursing homes	
	Coefficient	Significance [1]	Coefficient	Significance [1]
Independent variables				
Case mix, percent of patients with:				
Urinary tract infection	.118	.050	.233	<.001
Mental retardation and/or developmental disability	−.009	.501	−.016	.314
Dependent in toileting	.051	<.001	.052	<.001
Digestive system disorder	.058	.008	.044	.062
Facility-level quality Z-score by problem				
Subcutaneous and/or complicated skin condition	.435	.050	.291	.172
Immobility	.294	.265	.158	.583
Constant	6.747	<.001	6.806	<.001

[1]The level of significance is based on a two-tailed t-test.

NOTES: The dependent variable is nursing cost per patient day. The mean of the variable for 65 freestanding nursing homes is $10.14, and R^2 is .41 ($p < .001$). The mean of the variable for 74 freestanding and hospital-based nursing homes is $10.47, and R^2 is .45 ($p < .001$).

SOURCES: University of Colorado Health Sciences Center, Center for Health Services Research: Primary Study Data, 1980. Colorado Department of Social Services, Denver, Colorado: Medicaid Cost Reports, 1980.

Table 3 shows the relationships between case mix and quality combined and per-diem nursing cost. The underlying hypothesis in this case is that case mix and quality alone affect cost. The case-mix variables had similar coefficients and significance levels when included (Table 2), although the quality variables were generally statistically insignificant. (The exception was the skin condition variable in the freestanding sample, $p = .05$.) A similar pattern emerged for the total cost regressions (not shown). When the quality variables were included alone in cost regressions, the R^2 values were low, ranging from .06 to .13. These results are representative of those obtained in many regressions, using a variety of case mix and quality variables. They suggest that case mix is an important determinant of cost, but that the inclusion of quality variables does not add significantly to the explanation of cost variations.

Table 4

Illustrative regressions relating nursing cost per patient day to case mix, quality, and facility characteristics: 1980

Variables	65 freestanding nursing homes		74 freestanding and hospital-based nursing homes	
	Coefficient	Significance [1]	Coefficient	Significance [1]
Independent variables				
Case mix, percent of patients with:				
Urinary tract infection	.032	.562	.123	.009
Mental retardation and/or developmental disability	−.002	.871	−.012	.405
Dependent in toileting	.045	<.001	.058	<.001
Digestive system disorder	.024	.226	.021	.333
Facility-level quality Z-score by problem				
Subcutaneous and/or complicated skin condition	.305	.135	−.104	.621
Immobility	.223	.330	.283	.279
Facility characteristics				
Hospital-based	—	—	1.847	.014
Nonprofit	1.090	.005	1.178	.009
Percent Medicaid patients	−.006	.576	.009	.452
Percent nursing pool use	9.635	<.001	9.881	<.001
Constant	7.297	<.001	5.451	<.001

[1]The level of significance is based on a two-tailed t-test.

NOTES: The dependent variable is nursing cost per patient day. The mean of the variable for 65 freestanding nursing homes is $10.14, and R^2 is .58 ($p < .001$). The mean of the variable for 74 freestanding and hospital-based nursing homes is $10.47, and R^2 is .62 ($p < .001$).

SOURCES: University of Colorado Health Sciences Center, Center for Health Services Research: Primary Study Data, 1980. Colorado Department of Social Services, Denver, Colorado: Medicaid Cost Reports, 1980.

The facility characteristic variables were added to the regressions, with the results for nursing cost shown in Table 4 (similar results again emerged for the total cost regressions). This represents the complete model, including all factors hypothesized to affect costs. The addition of the facility characteristics increased the overall R^2 of the equations to about .60 and reduced the statistical significance of the individual case-mix variables, as is evident from a comparison of Tables 3 and 4. The toileting variable remained highly significant in both samples, but urinary tract infection remained highly significant in the sample of 74 facilities. All facility characteristics except the Medicaid variable were significant, and both quality variables were insignificant.

The results for the coefficient estimates in the various equations suggest that:

- Quality, as measured in these analyses, does not appear to be a major factor associated with cost variations.
- Case mix is related to cost, but case mix variations are also associated with certain facility characteristics.
- Hospital-based and nonprofit facilities have higher costs even after controlling for case mix and quality, so that only a portion of their higher costs can be attributed to case mix or quality differences.
- Nursing pool use is associated with higher costs.
- The percent of Medicaid patients is not related to cost once case mix and quality are taken into account.

Table 5 summarizes the proportion of variation in both nursing and total cost per patient day accounted for by the different sets of independent variables used in the earlier regressions. The case-mix regressions are presented in Table 2, in which the case mix variables alone yielded R^2 values of .33 and .43 for nursing cost, and .25 and .45 for total cost. When quality variables were added to case mix (Table 3), the R^2 values usually increased only slightly. Thus, the case-mix variables were found to be more strongly associated with cost than were the quality variables.

The addition of facility characteristics to the regressions (Table 4) increased the R^2 values to between .58 and .64. When facility characteristics were used alone in the cost regressions (not shown), they accounted for over 40 percent of the variation in both dependent variables for the two samples.

The results shown in Table 5 can be used as described to estimate the lower and upper bounds in the proportion of the cost variation accounted for by case mix and quality. An estimate of the lower bound is the amount by which the R^2 increases when case mix and quality variables are added to the facility characteristic variables. The upper bound estimate is the R^2 when case mix and quality are used alone in the regressions.

Thus, for the sample of 65 freestanding nursing homes, case mix and quality accounted for between 17-41 percent of the variation in nursing cost per patient day, and between 14-29 percent of the variation in total cost per patient day. These proportions were greater for the sample of 74 nursing homes and were almost the same for both dependent variables, ranging from 21-22 percent to 45-46 percent. As is also evident from this table and earlier results, the associations attributable to case mix and quality were primarily a result of case mix and not quality.

Table 5

Explanatory power (R^2) of case mix, quality, and facility characteristics in nursing and total cost per patient day regressions: 1980

Independent variables	Explanatory power (R^2)	
	65 freestanding nursing homes	74 freestanding and hospital-based nursing homes
Nursing cost per patient day		
Case mix	.33	.43
Case mix and quality	.41	.45
Case mix, quality, and facility characteristics	.58	.62
Facility characteristics alone	.41	.41
Total cost per patient day		
Case mix	.25	.45
Case mix and quality	.29	.46
Case mix, quality, and facility characteristics	.59	.64
Facility characteristics alone	.45	.42

NOTES: The case-mix variables are based on the percents of patients with urinary tract infection, mental retardation and/or developmental disabilities, dependent in toileting, and digestive system disorders. The quality variables are a facility-level quality Z-score, by problem for subcutaneous and/or complicated skin condition, and immobility. The facility characteristic variables are hospital-based, nonprofit, percent of Medicaid patients, and percent nursing pool use (N = 74 freestanding and hospital-based only).

SOURCES: University of Colorado Health Sciences Center, Center for Health Services Research: Primary Study Data, 1980.
Colorado Department of Social Services, Denver, Colorado: Medicaid Cost Reports, 1980.

The association of case mix and quality with facility characteristics

As discussed earlier, one objective of these analyses was to determine the extent that certain facility characteristics actually reflect differences in case mix and quality. The results show, by the weakening of some of the case-mix variables when facility characteristics were added to the cost regressions, an association between case mix and quality and facility characteristics. To examine this issue further, additional regressions were estimated. First, each facility characteristic was related separately to nursing and total cost (in both samples), and then were added to the case mix and quality variables. An indication of the degree that a particular facility characteristic served as a proxy for case mix and quality was then provided by the reduction in its regression coefficient or elasticity and in its statistical significance when the case mix and quality variables were added. The results of applying this procedure to the hospital-based or freestanding, forprofit or nonprofit, percent Medicaid patients, and percent nursing pool use variables are shown in Table 6 for nursing cost per patient day (the results for total cost were similar).

The addition of case mix reduced the positive coefficients of both the hospital-based and the nonprofit variables appreciably, whereas the addition of the quality variables had little effect. These results are consistent with the hypothesis that hospital-based and nonprofit nursing homes have a case mix of greater intensity than freestanding or forprofit nursing homes, and that such case mix differences account for at least part of the higher costs in these facilities. Quality differences, on the other hand, had no appreciable effect, and cannot be claimed (on the basis of these results) as reasons for higher costs in nonprofit or hospital-based nursing homes.

The Medicaid variable was statistically significant and negative for the sample of 65 freestanding facilities when used alone, but became statistically insignificant in both the nursing cost and total cost equations when the case mix variables were added. These results are consistent with other findings of a lower case-mix severity for Medicaid than for non-Medicaid patients (Shaughnessy, Schlenker, Yslas, et al., 1983). Thus, the negative elasticity of the Medicaid variable, when used alone in the sample of 65 facilities, appears to be largely the result of the less intense case mix associated with Medicaid patients in freestanding facilities. The percent of Medicaid patients was statistically insignificant in all regressions for the sample of 74 facilities.

The nursing pool variable was statistically significant and positive when used alone, and its elasticity and statistical significance lessened when case mix and quality variables were added. Although the changes were less dramatic than for the other facility characteristics, they suggest that nursing pool use may be associated with case mix and quality. Further exploration of this issue is desirable.

In summary, these results suggest that the hospital-based or freestanding, forprofit or nonprofit, and percent Medicaid variables act as partial proxy measures for case-mix differences. Quality, in contrast, appears to have a minimal association with these three facility characteristics.

Summary and conclusions

Major findings

In this study, case mix was found to be closely associated with cost. As much as 45 percent of the variation in cost per patient day was accounted for by variations in case mix among facilities. These results were not dependent on any one set of case-mix variables, because similar results emerged for several different sets of case-mix measures.

The relationship between quality and cost was considerably weaker. Although several quality variables were found to be significantly positively related to cost in preliminary analyses, representative quality variables by themselves accounted for only 10 percent of the variation in cost per patient day. When the quality variables were added to case-mix variables in the equations, they contributed little beyond case mix to the explanation of differences in cost per patient day among nursing homes. However, this study represents but an initial attempt to incorporate (process) quality measures aggregated from the patient-level to the facility level, these results cannot be considered definitive.

Table 6

Regression coefficients or elasticities of selected facility characteristics when case mix and quality variables are added to nursing cost per patient day regressions: 1980

Independent variables	65 freestanding[1] nursing homes			74 freestanding and hospital-based[1] nursing homes		
	Coefficient[1]	Significance[2]	R^2	Coefficient[1]	Significance[2]	R^2
Hospital-based						
Alone	—	—	—	2.682	<.001	.18
With case mix[3]	—	—	—	1.880	.003	.50
With case mix and quality[3,4]	—	—	—	—	—	—
Nonprofit						
Alone	1.173	.017	.09	1.820	<.001	.17
With case mix[3]	.875	.041	.38	1.260	.002	.50
With case mix and quality[3,4]	—	—	—	1.213	.003	.52
Percent Medicaid patients	Elasticity[1]			Elasticity[1]		
Alone	−.229	.007	.11	−.118	.237	.02
With case mix[3]	−.082	.347	.34	−.004	.961	.43
With case mix and quality[3,4]	—	—	—	—	—	—
Percent nursing pool use						
Alone	.050	<.001	.24	.033	.011	.09
With case mix[3]	.044	<.001	.47	.020	.070	.46
With case mix and quality[3]	.039	.001	.50	.017	.126	.47

[1]Regression coefficients are presented for the categorical (0/1) hospital-based and nonprofit variables; elasticities (at the mean) are presented for the continuous Medicaid and nursing pool variables.
[2]The level of significance is based on a two-tailed t-test.
[3]The case mix and quality variables used are those listed in Tables 2-5.
[4]Values for coefficients/elasticities, significance, or R^2s are not presented in those instances where the facility characteristics used already became insignificant ($p < .10$) due to case-mix variables or when addition of quality variables neither lowered the regression coefficient/elasticity nor lessened its significance.

SOURCES: University of Colorado Health Sciences Center, Center for Health Services Research: Primary Study Data, 1980. Colorado Department of Social Services, Denver, Colorado: Medicaid Cost Reports, 1980.

The major facility characteristics found to be related to cost were the attributes of hospital-based or freestanding, forprofit or nonprofit, the percent of Medicaid patients, and the percent of total nursing expenses for nursing pool services. When used alone, these variables accounted for approximately 40 percent of the variation in nursing home cost-per-patient day. However, they were associated with case-mix factors that also explained a substantial portion of the variation in cost. The associations between facility characteristics and cost were substantially weakened when case mix was taken into consideration. In fact, the case-mix variables totally accounted for the statistical significance of the relationship between the percent of Medicaid patients and cost. Thus, such facility characteristics appear to serve as partial proxy measures of case-mix differences among nursing homes.

Policy implications

The strong association between case mix and cost suggests that case mix should be incorporated into nursing home reimbursement beyond the simple classification of patients into level of care categories. This is especially important because of the increase in nursing home case-mix intensity expected as hospitals attempt to discharge acute care patients early, in response to Medicare's diagnosis-related group payment system and other pressures.

Because case mix is most directly related to patient care services (particularly nursing services), reimbursement systems should link case mix primarily to the patient care cost category. Several States, including Ohio, West Virginia, and Maryland, have separated costs into various categories for Medicaid reimbursement purposes and use different payment methodologies for each category. Patient care costs in these systems are reimbursed on the basis of assessments of patient service needs. Such systems should be analyzed further as they evolve. The same concepts appear to be appropriate in developing a prospective payment system for SNF care under Medicare (Stassen and Bishop, 1983).

To be viable, the incorporation of case mix into an operational nursing home reimbursement system requires that patient assessment activities be timely and efficient. This suggests that patient assessments to determine case mix for reimbursement purposes should be based on samples of patients in each facility and not be conducted more frequently than necessary.

The relatively weak association found between quality and cost suggests that quality (at least as measured in this study) cannot yet be closely linked to cost for reimbursement purposes. However, incentive payments tied to quality may still be reasonable as a

means to encourage nursing homes to provide higher quality care. Because of the weak association between quality and cost, such incentive payments could possibly lead to improvements in quality at relatively little increased public cost. The critical ingredient may be to provide the appropriate education, standards, and guidelines for nursing home administrators and patient care personnel. However, this premise should be tested on an experimental basis before actual implementation in an operational system. The recent experiment funded by the National Center for Health Services Research, which used outcome based incentive payments to nursing homes, should provide valuable information in this regard (Meiners, Heinemann, and Jones, 1982).

Further research

This study analyzed the relationships between cost, case mix, and quality at the facility level. However, patient-level analyses are also warranted, focusing on the degree to which case mix and quality of care affect the cost of treating individual patients. Fries and Cooney (1983) present an example of work underway in this area (i.e., the resource-utilization groups approach). In addition, this study was conducted in only one State. Although Colorado had many similarities to other States, further analyses of other States should be carried out to verify these findings.

Further research can also lead to important case mix and quality measurement improvements, which could ultimately be incorporated into operational programs. In the case-mix area, this study used measures of long-term care problems, ADL's, and (to a lesser extent) diagnoses as the main case-mix indicators. Several of the different case-mix measures yielded similar findings in the case-mix analyses. Representative "tracer" variables provided a reasonable picture of the association between case mix and cost. Composite measures that combine individual case-mix variables should therefore be further developed and compared to individual variables in terms of their ability to depict case-mix patterns and cost-case-mix associations.

In the quality area, refinements in the process quality approach are warranted. The possibility of outcome and composite measures should also be explored. The degree that process quality measures relate to the ultimate outcomes of long-term care could not be assessed from this study, and further research is needed on the relationship between process quality and the short- and long-run outcomes of patient care. It appears critical, however, to assess the strength of the association between quality and cost prior to considering a reimbursement approach that directly incorporates quality measures.

Acknowledgments

Major contributions to the study were made by Inez Yslas, Mary Polesovsky, Keith Brown, and Daniel Holub. Korbin Liu and Philip Cotterill of the Health Care Financing Administration provided valuable guidance and suggestions during the course of the overall study.

References

Birnbaum, H., Bishop, C., Lee, A. J., and Jensen, G.: Why do nursing home costs vary? The determinants of nursing home costs. *Med. Care* 19(11):1095-1107, Nov. 1981.

Bishop, C.E.: Nursing home cost studies and reimbursement issues. Vol. 1, No. 4. HCFA Pub. No. 03045. Office of Research, Demonstrations, and Statistics, Health Care Financing Administration. Washington. U.S. Government Printing Office, 1980.

Bishop, Christine E.: Nursing home cost studies: Comments and comparison. *Health Services Research*, 18(3):382-386, Fall 1983.

Fries, Brant E., and Cooney, Leo M.: *A Patient Classification System for Long-Term Care.* Grant No. 18-P-97757/1-01. Health Care Financing Grants and Contracts Report. HCFA Pub. No. 03179. Office of Research and Demonstrations, Health Care Financing Administration. Washington. U.S. Government Printing Office, Sept. 1984.

Meiners, Mark R., Heinemann, Gloria, D., and Jones, Brenda J.: "An evaluation of nursing home payments designed to encourage appropriate care for the chronically ill: Some preliminary findings." Presented at the American Economic Association Meeting, New York, Dec. 1982.

Palmer, Hans C., and Cotterill, Phillip G.: Studies of nursing home costs. In Vogel, R. J., and Palmer, H. C., eds. *Long-Term Care: Perspectives from Research and Demonstrations.* Health Care Financing Administration. Washington. U.S. Government Printing Office, 1983.

Polesovsky, Mary, and Shaughnessy, Peter: Colorado Nursing Home Study: Case Mix Measurement and Descriptive Profiles. Supporting Document B to Study Paper 5. Denver. University of Colorado, Health Sciences Center, Sept. 1982.

Scanlon, William, and Weissert, William: Nursing home cost function analysis: A critique. *Health Services Research*, 18(3):387-391, Fall 1983.

Schlenker, Robert, Shaughnessy, Peter, and Yslas, Inez: The effect of case mix and quality on cost differences between hospital-based and freestanding nursing homes. *Inquiry*, 20(4):361-368, Winter 1983.

Shaughnessy, Peter W., Breed, Linda D., and Landes, David P.: Assessing the quality of care provided in rural swing-bed hospitals. *Quality Review Bulletin*, 8(5):12-20, May 1982.

Shaughnessy, Peter, Schlenker, Robert, and Yslas, Inez: Case Mix, Quality, and Cost: Major Findings and Implications of the Colorado Nursing Home Study. Study Paper 5. Denver. University of Colorado, Health Sciences Center, Feb. 1983.

Stassen, Margaret, and Bishop, Christine E.: *Incorporating case mix in prospective reimbursement for SNF under Medicare: Critical review of relevant research.* Grant No. 18-P-97038/1-05. Background report prepared for the Health Care Financing Administration. Waltham, Mass. Center for Health Policy Analysis and Research, Brandeis University, March 1983 (revised).

Yslas, Inez, Brown, Keith, and Shaughnessy, Peter: Colorado Nursing Home Study: Quality Measurement and Descriptive Profiles. Supporting Document C to Study Paper 5. Denver, Colo. University of Colorado Health Sciences Center, Sept. 1982.