Good quality care increases hospital profits under prospective payment

by David C. Hsia and Cathaleen A. Ahern

This study shows that, contrary to popular belief, the prospective payment system discourages skimping on medically indicated care. The quality of care on a nationally representative sample of Medicare discharges underwent judgmental review using implicit criteria. The reviewing physicians identified hospitalizations that omitted medically indicated services and diagnoses overlooked because of this skimping. After deduction for the cost of the omitted services and probability of negative diagnostic tests, good quality care would have increased hospital profits a significant 7.9 percent. As the specificity of diagnosis and intensity of treatment increase, the DRG payment rises faster than the cost of providing medically indicated services.

Background

Since October 1, 1983, Medicare has used a prospective payment system (PPS) to pay hospitals for inpatient care, as required by the Social Security Amendments of 1983. Each discharge's diagnoses and procedures "group" it to one of 477 diagnosis-related groups (DRGs). The Health Care Financing Administration (HCFA) pays the hospital a fixed amount representing the average cost for that DRG's discharges (*Code of Federal Regulations*, 1988). The hospital retains surpluses from discharges that cost it less than the DRG payment, and suffers losses on discharges that cost it more. The prospective payment system gives hospitals a financial incentive to reduce the unnecessary services associated with cost-based, retrospective payment.

Congress and other observers fear that prospective payment could induce hospitals to reduce necessary services (or access) (Newhouse, 1989), in addition to unnecessary services (U.S. House of Representatives, 1985; U.S. Senate, 1985; U.S. Senate, 1986b; U.S. House of Representatives, 1986a; U.S. House of Representatives, 1986b; U.S. Senate, 1986a; Prospective Payment Assessment Commission, 1990). The hospital's need to generate financial surpluses may pressure its attending physicians to omit medically indicated tests and therapies. The institution reduces costs, while continuing to receive a DRG payment intended to cover all necessary services. This "skimping" on the quality of care may place the patient's health at risk. If undertreatment causes the patient subsequently to need readmission, the poor quality confers a second payment upon the hospital (assuming no peer review organization intervention) (Munoz et al., 1990). This article tests the validity of the commonly held belief that prospective payment rewards such skimping.

At discharge, the attending physician writes each Medicare inpatient's diagnoses and procedures on a federally required attestation statement. The hospital's medical records department translates these narrative diagnoses and procedures into the *International Classification of Diseases 9th Revision, Clinical Modification* (ICD-9-CM) numeric codes (Public Health Service and Health Care Financing Administration, 1980). The fiscal intermediary groups the ICD-9-CM codes to the proper DRG, converts the corresponding relative weight to a dollar amount (with certain minor adjustments for hospital-specific factors), and pays the hospital (Averill et al., 1986).

Omission of medically indicated procedures could cause diagnostic uncertainty and may therefore produce vague ICD-9-CM codes (e.g., the classic weak, tired, and dizzy). For some discharges, good quality care (which includes an adequate workup) increases diagnostic specificity and changes the DRG assignment. When more complete workup reassigns the discharge to a DRG paid at a higher rate (adjusted for the cost of additional services and the possibility of the workup ruling out the suspected pathology), PPS rewards the hospital for good quality care. Omission of some therapeutic services also changes the DRG and therefore payment. However, omission of other services causes no change in payment, creating an economic disincentive to their delivery.

The following examples illustrate how PPS rewards (or fails to reward) quality care in the form of an adequate diagnostic workup. Suppose an elderly patient presents with sudden onset of severe chest pain, a non-specific symptom that groups to DRG 143. During Federal fiscal year 1985, PPS would have paid an urban hospital $2,013 for this admission (*Federal Register*, 1984). If the patient then receives an electrocardiogram that establishes ischemia, the discharge would group to DRG 140. Its $2,230 payment for angina exceeds the DRG 143 payment by $217. For 1985, Medicare Part B pays $28 for an electrocardiogram, an estimate of its average cost. Subtracting this service's cost, the hospital still would gain $189 "at the margin" (Samuelson, 1985). The hospital therefore has a financial incentive not to skimp on quality by omitting this medically indicated service (Table 1).

Further suppose that upon confirmation of angina, the attending physician orders a routine cardiac screen of three serial electrocardiograms ($28 each in 1985) and three sets of cardiac enzymes ($19 each). These tests confirm a myocardial infarction, grouping the hospital stay to DRG 122 (Lee, 1986). Its $4,032 payment would confer $1,802 more revenue than DRG 140. Deducting

Reprint requests: David C. Hsia, M.D., Department of Health and Human Services, Office of Inspector General, 1-D-16 Oak Meadows Building, 6340 Security Boulevard, Baltimore, Maryland 21207.

Table 1

Examples of the effect of increasing diagnostic specificity on Medicare prospective payment: Fiscal year 1985

DRG	Medical condition	Relative weight	Payment	Payment increase
Example 1				
143	Chest pain	0.6743	$2,013	—
140	Angina pectoris	0.7470	2,230	$217
122	Myocardial infarction	1.3509	4,032	1,802
109	Myocardial infarction with bypass	3.6574	10,917	6,885
Example 2				
464	Weakness	0.7246	2,163	—
12	Hemiplegia	1.1020	3,289	1,126
14	Non-specific cerebrovascular accident	1.3386	3,996	707
14	Intracerebral bleeding or cerebral arterial occlusion	1.3386	3,996	0
Example 3				
411	History of malignancy without endoscopy	0.7146	2,133	—
412	History of malignancy with endoscopy	0.3365	1,004	1,129

NOTE: DRG is diagnosis-related group.

SOURCE: (*Federal Register*, 1984).

the Part B payment of $141, an approximation of procedure costs, still would leave the hospital with an increased profit.

Finally, a major cardiac workup including stress test ($109), nuclear scan ($169), echocardiography ($128), catheterization ($518), and bypass ($3,897) could push the discharge into DRG 109. The $10,917 payment for a major cardiac procedure exceeds the DRG 122 payment by $6,885. After deducting the cost of these procedures and the additional length of stay, the hospital should still profit more than for DRG 122.

However, PPS does not always reward good quality care's diagnostic specificity with higher payment. Where the physician suspects a stroke, the increasingly specific diagnoses of weakness (DRG 464), hemiplegia (DRG 12), and cerebrovascular accident (DRG 14) carry rising payments. However, computer tomography ($97) to distinguish between hemorrhagic and ischemic disease does not change the DRG. This test, commonly used to manage the cerebrovascular accident, would confer no additional payment on the hospital.

Indeed, provision of medically indicated services may even reduce payment in some cases. A patient with a history of colon carcinoma generally needs endoscopic followup at some point. However, in 1985, DRG 412 would have paid one-half as much for malignancy after care with endoscopy, as DRG 411's payment for aftercare without endoscopy (Brooks, 1984). Presumably, the population of patients grouping to DRG 411 still consumes more resources than the DRG 412 population (e.g., because of more advanced disease or more expensive treatment). Alternately, did the lower DRG 412 payment deter endoscopy, confusing cause and effect?

Returning to the first example, not all electrocardiograms confirm chest pain to be angina (Fisch, 1988). The proportion of positive electrocardiograms varies depending on different populations' health status and physicians' test-ordering patterns, but averages about 50 percent. In performing the service, the hospital would have a 0.5 probability of receiving $2,202 ($2,230 payment for DRG 140 minus $28 test costs) and a 0.5 probability of receiving $1,985 ($2,013 payment for DRG 143 minus $28 test costs). Statistically, the hospital has an expectation of $2,094. If it omits the study, it has a 1.0 probability of receiving $2,013 for DRG 143 and incurs no extra diagnostic costs. Based on these limited considerations, the economically rational hospital should perform electrocardiograms on patients presenting with appropriate chest pain symptoms because on average, it would retain net revenue of $81 after deducting the extra cost and discounting for negative tests. In this case, good quality care increases profits (Table 2).

Some of the other examples entail slightly more elaborate expectation calculations. About 29 percent of cardiac workups lead to bypass surgery, and a workup with surgery costs more than a workup without surgery. The hospital would therefore have a 0.29 probability of receiving $6,096 ($10,917 for DRG 109 minus $4,821 procedure costs) and a 0.71 probability of receiving $3,108 ($4,032 for DRG 122 minus $924 costs), for an expectation of $3,975. Not doing the medically indicated workup carries a 1.0 probability of $4,032. In this situation, providing proper care would penalize the hospital $57.

Finally, DRG 412 depends on performing endoscopy, rather than on the results of endoscopy. Since the physician largely controls the decision to do the endoscopy, it has a probability of 100 percent. That 6 percent of endoscopies yield positive findings has no bearing on the expectation calculation (Matek et al., 1985). This evaluation assumes that issues of diagnostic reliability (i.e., false positives and negatives) have little effect on payment in comparison to the other variables.

Under this analysis, a DRG sequence of increasing diagnostic specificity may increase profits. Conversely, skimping on medically indicated services maximizes hospital profits in other DRG sequences. This article gauges which trend has the greater economic effect. Overall, does prospective payment reward or penalize good quality care? Obviously, multiple factors influence medical workup and treatment decisions. In this article, we consider only the economic

Table 2

Marginal analysis of the effect of increasing diagnostic specificity on Medicare prospective payment: Fiscal year 1985

Diagnostic data	DRG 143 to DRG 140 change Electrocardiogram	DRG 140 to DRG 122 change Cardiac enzymes	DRG 122 to DRG 109 change Cardiac workup	DRG 464 to DRG 12 change History	DRG 12 to DRG 14 change Physical examination	DRG 14 to DRG 14 change CAT scan	DRG 411 to DRG 412 change Endoscopy
Test positive							
Probability	[1,2]0.50	[3]0.58	[4]0.29	[5]0.83	[6]0.30	[7]0.20	[8]1.00
Payment	2,230	4,032	10,917	3,289	3,996	3,996	1,004
Test cost	−28	−141	−4,821	−30	−35	−97	−57
Expectation[9]	1,101	2,257	1,768	2,705	1,188	780	947
Test negative							
Probability	0.50	0.42	0.71	0.17	0.70	0.80	0.00
Payment	2,013	2,230	4,032	2,163	3,289	3,996	2,133
Test cost	−28	−141	−924	−30	−35	−97	−57
Expectation	993	877	2,207	363	2,278	3,119	0
Total expectation[10]	2,094	3,134	3,975	3,068	3,466	3,899	947
Payment without test	2,013	2,230	4,032	2,163	3,289	3,996	2,133
Difference	81	904	−57	905	177	−97	−1,186

[1]Behar, S., Schor, S., Kariv, L., et al.: Evaluation of electrocardiogram in emergency room as a decision-making tool. *Chest* 71:486, 1977.
[2]McGuinness, J.B., Begg, T.B., and Semple, T.: First electrocardiogram in recent myocardial infarction. *British Medical Journal* 2:449, 1976.
[3]Sobel, B.E., and Shell, W.E.: Serum enzyme determinations in the diagnosis and assessment of myocardial infarction. *Circulation* 54:471-482, 1972.
[4]Proudfit, W.L., Welch, C.C., Siqueira, C., et al.: Prognosis of 1,000 young women studied by coronary angiography. *Circulation* 64:1185-1190, 1981.
[5]Fischer, C.M.: Development of the clinical picture in 125 cases of cerebral thrombosis. In Adams, R.D., and Victor, M., eds. *Principles of neurology*. 4th ed. New York. McGraw-Hill, 1989.
[6]Mills, M.L., Russo, L.S., Vines, F.S., and Ross, B.A.: High-yield criteria for urgent cranial computed tomography scans. *Annual Emergency Medicine*, 15:1167-1172, 1986.
[7]Mohr, J.P., Kase, C.S., and Adams, R.D.: Cerebrovascular diseases. In Petersdorf, R.G., et. al. *Harrison's principle of internal medicine*. 10th ed. New York. McGraw-Hill, 1983.
[8]Physician-controlled.
[9]Probability x (payment − test cost).
[10]Expectation test negative plus expectation test positive.

NOTES: DRG is diagnosis-related group. CAT is computerized axial tomography.

SOURCE: (Health Care Financing Administration, 1984).

consequences of poor quality. Non-economic behaviors such as altruism, professionalism, or fear of malpractice litigation are not measured in this study.

Methods

The National DRG Validation Study employed a stratified two-stage sample design based on hospitals and discharges (Delaney, 1987). In the first stage, the Office of Inspector General (OIG), U.S. Department of Health and Human Services, used simple random sampling without replacement to select 80 hospitals from each of three bed-size strata: hospitals with fewer than 100 beds, 100 to 299 beds, and 300 beds or more. If quality of care varied by hospital size, as expected, this stratification maximized the statistical information. The design excluded specialty institutions (e.g., pediatric, rehabilitation, and psychiatric hospitals), facilities in States not using Federal prospective payment during the period studied (i.e., New York, New Jersey, Massachusetts, and Maryland), and hospitals not contributing data to the calculation of the initial relative weights assigned to DRGs. One sample hospital terminated its Medicare eligibility between the study period and actual collection of medical records, leaving a first-stage sample of 239 from a population of 4,913 acute care hospitals (Table 3).

The second stage used systemic random sampling to select up to 30 Medicare discharges (including patients who transferred to other hospitals or died) from each of the 239 hospitals for the first half of Federal fiscal year 1985: October 1, 1984 to March 31, 1985. If the hospital had fewer than 30 such discharges during this period, the sample selected all available discharges. OIG then requested a complete copy of each of the 7,076 selected medical records. With careful followup and selective use of subpoenas, it ultimately obtained 7,050 charts (99.6 percent) representing 6,900 different patients. Comparison of the records' demographic characteristics demonstrated that the sample accurately represented the population of all Medicare beneficiaries in PPS jurisdictions (Hsia, 1988).

OIG contracted with the Baxter-Health Data Institute of Lexington, Massachusetts for registered records analysts and accredited records technicians to perform a blinded reabstraction of the ICD-9-CM disease codes supported by the chart (Ahern, 1988). In addition, specially trained registered nurses screened each record for quality of care using the Appropriateness Evaluation Protocol, a chart audit instrument of

Table 3

Sampling frame, by hospital bed size: Fiscal year 1985

Data category	Total	Bed size		
		Fewer than 100	100 to 299	300 or more
Hospitals				
Medicare	4,913	2,536	1,603	774
Sample	239	79	80	80
Sampling fraction (percent)	4.9	3.1	5.0	10.3
Discharges				
Medicare	8,277,000	1,522,000	3,105,000	3,649,000
Population	222,396	18,199	59,481	144,716
Sample	7,050	2,276	2,388	2,386
Sampling fraction (percent)	3.2	12.5	4.0	1.6

SOURCE: (Hsia et al., 1988).

proven utility (Siu, 1986). Cases failing quality screening underwent judgmental review by contractor physicians using implicit criteria to identify care clearly failing to meet "professionally recognized standards," the statutory criterion for peer review organizations (United States Code, 1989). The reviewing physicians had instructions to ignore borderline problems or legitimate differences in medical judgment as to appropriate management (Delaney and Hsia, 1987).

Upon identification of a discharge with one or more irrefutable quality problems (e.g., non-workup of life-threatening and potentially reversible symptoms, non-delivery of essential medication), the physician reviewer dictated a narrative summary describing the nature of the deficiency and citing supporting evidence from the chart. The reviewers had extensive chart audit experience, board certification in pertinent clinical specialties, and recent patient-care responsibility. Appropriate specialists reviewed records presenting specialty care issues. Physician panels decided difficult cases. Reliability checks demonstrated no significant misclassifications (agreement = 0.994, Kappa = 0.963, Z = 20.8) (Cohen, 1960).

Physicians identified four types of poor quality:

- Omission of medically indicated services (skimping).
- Provision of unnecessary services.
- Complication to indicated services (e.g., postoperative infection).
- Other.

Discharges classified as having only unnecessary services, complications, or "other" did not undergo further review because PPS provides no economic incentives promoting such behaviors. The physicians then classified the skimping discharges by type of service omitted:

- History and physical examination.
- Laboratory test (e.g., blood glucose).
- Radiology or non-invasive imaging (e.g., ultrasound).
- Other diagnostic procedure (e.g., colonoscopy, biopsy, or other invasive procedure).
- Therapy (e.g., medication, surgery).
- Other.

Finally, they identified whether the omitted services could have caused a change in ICD-9-CM codes. They selected revised diagnosis and therapy codes without knowing how these changes would affect selection DRG classification and its payment consequences. The classifiers anecdotally observed that they had no difficulty choosing the revised codes. Reliability checks disclosed no significant disagreements about revised diagnoses (agreement = 0.973, Kappa = 0.941, Z = 19.0).

Medicare-approved grouper software processed the resulting ICD-9-CM codes to determine any new DRGs resulting from addition of revised diagnoses identified by reviewers, and to assign relative weights and corresponding dollar payment. Medicare data files supplied the average Part B payment for each omitted service. These estimates of procedure cost did not warrant adjustment for increased length of stay because the omitted services proved to be minor procedures not prolonging hospitalization. For this reason, the methodology also did not adjust for the probability of complications to the omitted services. Medline literature searches provided information about the probabilities of each diagnostic test's yielding a positive result. Spreadsheet software calculated the expectation, average change, and total change resulting from the independent variables.

A sensitivity analysis identified the variables that had the greatest influence on the final result (Stokey and Zeckhauser, 1978; Mason, 1987; Leamer, 1985). Note that despite a similar nomenclature, economic and public policy "sensitivity" bears no relationship to epidemiological "sensitivity," the percentage of positive tests among the populations of individuals with the disease (Budnick, 1987). Rather, where a projected result depended on accurate measurement of a sequence of related, independent variables, sensitivity analysis successively substituted probable high and low values for each variable for the usual point estimate (Poister, 1978). This technique produced a range (or interval) of probable results in place of the usual single result.

Dividing the output interval by each input variable's interval calculated the β statistic (beta) as a derivative. The βs quantified the outcome's sensitivity to changes in the input variables. High β variables strongly affected the outcome and therefore warranted maximum accuracy. Low β variables had little impact on study results and therefore warranted only limited

policy attention. Some other sensitivity analyses preferred to calculate the β as a ratio, rather than as a derivative.

In this case, six independent variables—rate of poor quality care (of all four types), proportion of poor quality because of skimping, proportion of skimping discharges with revised DRGs, net dollar payment for revised DRGs, test costs, and proportion of positive tests—determined the dependent variable: dollar change in payment. The sensitivity analysis used the 95-percent confidence interval for each of these input variables as its estimated range. As a reality check, a second sensitivity analysis used numbers reported in the previous literature.

Results

Quality

Of the 7,050 sample discharges, reviewing physicians identified 464 (5.5 percent on strata weighting by 1985 Medicare discharges) as failing to meet professionally recognized standards for quality of care (Admire et al., 1989). Smaller hospitals had a significantly higher rate of quality problems than larger hospitals (Chi-square 120.5, 2 df, $P < 0.0001$). The good and poor quality subsamples did not differ in sex distribution (Mantel-Haenszel 0.5, 1 df, $P = 0.47$) when controlling for hospital size, but older inpatients suffered a significantly higher rate of poor quality (Mantel-Haenszel 16.2, 1 df, $P = 0.001$) (Mantel and Haenszel, 1959) (Table 4).

Skimping

Of the 464 patients receiving poor quality care, 87.9 percent experienced at least one instance of omitting medically necessary services. The distribution of skimping by hospital size did not significantly differ from the bed size distribution for poor quality care of all types (Chi-square 0.6, 2 df, $P = 0.76$). Inadequate diagnosis comprised 80.0 percent of the 758 individual instances of skimping, with undertreatment making up the balance (Table 5).

Revised diagnosis-related groups

Among the 408 discharges with skimping, proper care would have caused 63.7 percent to group to other DRGs. The proportion did not significantly differ by hospital size (Chi-square 2.3, 2 df, $P = 0.32$). Of these 260 discharges with revised diagnoses, 79.2 percent initially would have grouped to higher weighted DRGs. However, 11.2 percent of these higher weighted discharges would have become unprofitable upon deduction for the probability of negative tests and testing costs. Good quality of care therefore would have increased the hospital's profits in 44.9 percent of the 408 skimping discharges (Table 6).

Payment

For all 408 skimping discharges as a group, delivery of medically indicated services still would have had a beneficial effect on hospital profits. The three groups (revised DRG higher, lower, or same) had similar estimated testing costs (averaging $56) and estimated probabilities of positive tests (averaging 50.6 percent). However, the first group's higher frequency and larger payment difference would more than offset the other, unprofitable groups. On average, good quality care would have increased profits by a significant 7.9 percent or $147 per discharge (t-test 6.3, 406 df, $P < 0.05$) (Table 7).

Table 4

Quality of care, by patient demography: Fiscal year 1985

	Total		Bed size		
Discharges	Number	Percent of total	Fewer than 100	100 to 299	300 or more
Total	7,050	100.0	2,276	2,388	2,386
Good quality care					
Under 65 years	678	9.6	128	248	302
65-74 years	2,786	39.5	791	930	1,065
75-84 years	2,162	30.7	704	781	677
85 years or over	960	13.6	394	307	259
Male	3,007	42.7	895	1,022	1,100
Female	3,579	50.8	1,132	1,244	1,203
Subtotal	6,586	93.4	2,017	2,266	2,303
Poor quality care					
Under 65 years	41	0.6	20	12	9
65-74 years	150	2.1	77	42	31
75-84 years	148	2.1	84	40	24
85 years or over	125	1.8	78	28	19
Male	216	3.1	111	61	44
Female	248	3.5	148	61	39
Subtotal	464	6.6	259	122	83

SOURCE: (Admire et al., 1989).

Table 5

Poor quality of care, by type: Fiscal year 1985

Data category	Total	Bed size		
		Fewer than 100	100 to 299	300 or more
Type of poor quality				
Skimping	408	238	101	69
Unnecessary service	47	20	20	7
Complication	131	69	29	33
Other	47	30	9	8
Total poor quality	464	259	122	83
Type of skimping				
History and physical	92	60	22	10
Laboratory tests	205	123	53	29
Radiology	150	89	38	23
Other diagnostic	159	107	33	19
Therapy	143	73	42	28
Other	9	5	1	3
Total skimping	408	238	101	69

NOTE: Columns do not add to total because a discharge may appear in multiple rows.

SOURCE: *Office of Inspector General: Data from the National DRG Validation Study.*

Table 6

Skimping, by revised diagnosis-related groups (DRGs): Fiscal year 1985

Data category	Total	Percent of total	Bed size		
			Fewer than 100	100 to 299	300 or more
Total	408	100.0	238	101	69
DRG revised	260	63.7	158	58	44
DRG not revised	148	36.3	80	43	25
Revised DRG					
Total	260	100.0	158	58	44
Higher weight	206	79.2	132	41	33
Lower weight	54	20.8	26	17	11
Higher weight DRGs					
Total	206	100.0	132	41	33
Payment still higher	183	88.8	120	36	27
Payment not still higher	23	11.2	12	5	6

SOURCE: Office of Inspector General: Data from the National DRG Validation Study.

Sensitivity analysis

For each input variable, hospital profits increased throughout their confidence intervals. While measurement variation could have slightly increased or decreased profitability, it could not have caused unprofitability. The proportion of skimping, revised DRGs, and positive tests had low βs, so that large changes in their values had little effect on payment. Rate of poor quality care, payment changes, and test costs had high βs, meaning that changes in their values strongly influenced the final results. Fortunately, the latter two variables depended exclusively upon objective, reproducible billing records, and published measurements provided a comparison for the rate of poor quality care (Table 8).

A sensitivity analysis also tested previously reported values for the study's variables, where available. Input to the model, these ranges confirmed that good quality care consistently increased profits. They also corroborated the relative magnitudes of the βs (Table 9).

Discussion

Quality

The previous literature infers the quality of care from the properties of structure, process, or outcome (Donabedian, 1966; Donabedian, 1982). Structure refers to inherent provider characteristics, e.g., does the attending physician have a current medical license? (Donabedian, 1968). Process describes provider actions, e.g., does the physician order a medically indicated test? Outcome pertains to the effect of provider actions, e.g., mortality, morbidity (Donabedian, 1980). The literature vigorously debates the merits of each type of measure, generally proposing outcome measures as a theoretical ideal and process measures as the practical reality (American Medical Association Council on Medical Science, 1986). Peer review organizations and malpractice litigation both utilize process measures as established by implicit, judgmental techniques.

The present study uses implicit process measures to classify inpatient care as either good or poor. Process measures have higher variance than structure or outcome measures, but most closely approximate the

Table 7

Financial effects of not skimping, by cost component: Fiscal year 1985

Revised-DRG data	Higher	Lower	Same	Average
Total change[1]	$107,590	-$40,176	-$7,401	$60,012
Discharges	206	54	148	408
Payment	$2,386	$3,866	$3,146	$2,858
Revised payment	$3,686	$2,562	$3,146	$3,341
Test cost	$62	$46	$50	$56
Positive tests (percent)	45.0	53.5	57.2	50.6
Expectation[2]	$2,909	$3,122	$3,096	$3,005
Average change[3]	$523	-$744	-$50	$147
Average change (percent)[4]	21.9	-19.2	-1.6	7.9

[1]Average change x discharges.
[2](positive test x (revised payment - test cost)) + ((1 - positive tests) x (payment - test cost)).
[3]Expectation - payment.
[4]Average change/payment.

NOTE: DRG is diagnosis-related group.

SOURCE: Office of Inspector General: Data from the National DRG Validation Study.

Table 8

Sensitivity analysis of 95-percent confidence interval: Fiscal year 1985

Variable	Point estimate	95-percent CI	Payment change	β[1]
		Percent		
Poor quality of care	[2]6.6	6.0-7.2	$54,723-65,667	0.912
Skimping	[3]87.9	85.0-90.9	58,033-62,062	0.068
Revised DRGs	[4]63.7	59.1-68.4	54,206-65,945	0.126
Payment change	[5]7.9	5.7-11.3	31,906-99,072	1.199
Test cost	[6]2.0	1.6-2.3	55,574-64,451	1.268
Positive tests	[7]50.6	46.3-54.8	57,147-62,878	0.067

[1]Percent change required for $10,000 payment increase: payment change / (95-percent CI range x $10,000).
[2]Table 4, line 2.
[3]Table 4, line 4.
[4]Table 5, line 1.
[5]Table 6, line 8.
[6]Table 6, line 4 / table 6, line 2 = percent of payment.
[7]Table 6, line 5.

NOTE: CI is confidence interval. DRG is diagnosis-related group.

SOURCE: Office of Inspector General: Data from the National DRG Validation Study.

Table 9

Sensitivity analysis of previous literature's range

Variable	Percent	Payment change	β
Poor quality of care	[1]4.7-[2,3]7.2	$42,866-65,667	0.912
Test cost	[4]0.8-[5]2.5	54,110-72,867	1.103
Positive tests	[6]6.0-[7]58.5	4,149-92,593	0.168

[1](Mills, 1977).
[2](Piltch, 1988).
[3](Health Care Financing Administration, Health Standards and Quality Bureau, 1986).
[4]Wetzel, A.M., and Kirz, D.S.: Routine hepatitis screening in adolescent pregnancies—Is it cost effective? *American Journal of Obstetrics and Gynecology* 156:166-169, 1987.
[5]Weincek, R.G., Weaver, D.W., Bouwman, D.L., and Sachs R.J.: Usefulness of selective preoperative chest X-ray films. *American Surgery* 53:396-398, 1987.
[6]Frye, E.C., Hubbell, F.A., Akin, B.V., and Rucker, L.: Usefulness of routine complete blood cell counts on a general medical service. *Journal of General Internal Medicine* 2:373-376, 1987.
[7]Thorson, A.G., Christensen, M.A., and Davis, S.J.: The role of colonoscopy in the assessment of patients with colorectal cancer. *Diseases of the colon and rectum* 29:306-311, 1986.

SOURCE: Office of Inspector General: Data from the National DRG Validation Study.

physician's reasoning process (Brook, 1974). The National DRG Validation Study's finding of 5.5-percent poor quality care falls in between the 4.7 (Mills, 1977) and 7.2 percent reported in previous studies (Health Care Financing Administration, 1986; Piltch, 1988). These similar rates of poor quality care indicate the judgmental methodology to be reproducible and valid. Upon sensitivity analysis, the previous literature's proportions do not affect this article's conclusion that good quality increases hospital profits from Medicare.

Skimping

Confirming congressional concerns, skimping on services proves to be the most common type of poor quality care. Indeed, previous literature discusses the use of skimping as a profit maximizing strategy (Hobler et al., 1985; Stern and Epstein, 1985; Kahkonen, 1985).

Unfortunately, no studies establish the national proportion of skimping prior to 1983. Accordingly, this article cannot prove that skimping increases under prospective payment. In any event, because of its low β, variation in the rate of skimping has little effect on prospective payment profits.

Revised diagnosis-related groups

Selection of revised diagnoses introduces relatively little error. The classifiers have high inter-rater agreement. This finding parallels anecdotal observations that hospital-based peer review committees have little difficulty deciding what diagnoses their colleagues should have worked up in a particular clinical situation. The sensitivity analysis identifies this variable as having a low β. It can change Medicare profitability only slightly.

Payment

Change in payment naturally has a much greater effect on the results. However, determining payment offers little opportunity for error. The computerized grouper automatically converts the ICD-9-CM codes into a payment. Except for data entry errors, input of a given combination of ICD-9-CM codes always produces the same payment as its output. This variable's high reproducibility therefore limits any misclassification effects from its high β.

Test cost

Part B payment serves only as a crude approximation of actual test costs. Medicare receives some criticism for its payments being low in comparison with those of private insurers (Firshein, 1986). Physicians and providers assert that for selected services the Federal compensation barely covers their costs. HCFA in turn vigorously defends its methodology for setting payments (Price, 1989). Practically, beneficiaries can obtain virtually all Part B services at the Medicare price in essentially all geographic areas (Garrison, 1986). Either the payment covers costs or the provider behaves non-economically, e.g., out of a sense of professional duty to a long-time patient (Goodwin and Dolan, 1985). In any event, the sensitivity analysis demonstrates that procedure costs can have little impact on the profitability of good quality care.

Positive tests

The proportion of positive results necessarily varies from test to test. In addition, populations differ in their prevalences (e.g., cardiac screening of asymptomatics at a college versus a nursing home), and physicians differ in the clinical thresholds that trigger particular workups (e.g., computerized tomography scan for headache at an underutilized community hospital versus an overloaded public hospital) (Thompson and Krushat, 1989). This article uses secondary sources for its estimates of the proportion of tests having positive results. Had these sources used different study populations or selection methodologies, they could have reported higher or lower test yields. Fortunately, the sensitivity analysis indicates that this variable has a lesser effect on profits than do other variables. Substitution of ranges from the previous literature confirms the β.

The foregoing results refute the commonly held belief that prospective payment encourages skimping on the quality of care. Economic theory states that where revenue remains fixed, the firm should cut costs to maximize profits. In actuality, the DRG system does not necessarily fix payment. Many of its "major diagnostic categories" contain sequences of DRGs, whose diagnosis and treatment entail increasing levels of service, balanced by rising payment. Overall, not skimping on quality produces significantly higher profits despite the addition of test costs and allowance for negative tests. Accordingly, conventional wisdom is wrong in this instance.

Acknowledgments

The authors are indebted to W. Mark Krushat, Linda M. Moscoe, Brian P. Ritchie, and Barry L. Steeley of the Office of Inspector General; John R. Boyle and Walter P. Weaver of the Office of the Assistant Secretary of Management and Budget; Annette M. Delaney, Laurie H. Moore, Joseph S. Yarmus, and their staff at the Baxter-Health Data Institute; and the anonymous reviewers for their technical assistance with this project.

References

Admire, K.L., Czajka, L., Carlson, R.E., et al.: National DRG validation study—Quality of patient care in hospitals. Control No. OAI-09-88-00870. Office of Inspector General, U.S. Department of Health and Human Services. Seattle, WA., 1989.

Ahern, C.A.: Special report on coding accuracy. Control No. OAI-12-88-01010. Office of Inspector General, U.S. Department of Health and Human Services. Washington, D.C. 1988.

American Medical Association Council on Medical Science: Quality of care. *Journal of the American Medical Association* 256:1032-1034, 1986.

Averill, R.F., Mullin, R.L., Giardi, P.A., and Elia, E.D.: *Diagnosis related groups, third revision—Definitions manual.* New Haven, CT. Health Systems International, 1986.

Brook, R.H.: A skeptic looks at peer review. *Prism* 2(10):29-32, Oct. 1974.

Brooks, P.E., and Fagan, A.B.: DRG creep and miscoding under the prospective payment system. Office of Inspector General, U.S. Department of Health and Human Services. Washington, D.C. 1984.

Budnick, L.D.: Statistics. In Cassens, B.J., *Preventive Medicine and Public Health.* Media, PA. Harwal Publishing, 1987.

Code of Federal Regulations: Public Health. Title 42, Part 412. Prospective payment for inpatient hospital services. Office of the Federal Register, National Archives and Records Administration, Washington. U.S. Government Printing Office, Oct. 1, 1988.

Cohen, J.: A coefficient of agreement for nominal scales. *Educational Psychology Measurement* 20:37-46, 1960.

Delaney, A.M., and Hsia, D.C., eds: *National DRG Validation Study—Final report*. Lexington, MA. Baxter-Health Data Institute, 1987.

Donabedian, A.: Evaluating the quality of medical care. *The Milbank Memorial Fund Quarterly*, Health and Society 44 (Part 2):166-206, July 1966.

Donabedian, A.: Promoting quality through evaluating the process of patient care. *Medical Care* 6:181, 1968.

Donabedian, A.: *Definition of quality and approaches to its assessment*. Ann Arbor, MI. Health Administration Press, 1980.

Donabedian, A., Wheeler, J.R., and Wyszewianski, L.: Quality, cost, and health—An integrative model. *Medical Care* 20(10):975-92, Oct.1982.

Federal Register: Medicare Program: Prospective payment update. Vol. 49, No. 171, 34723-34789. Office of the Federal Register, National Archives and Records Administration. Washington. U.S. Government Printing Office, Aug. 31, 1984.

Firshein, J.: Providers call '87 PPS increase unacceptable. *Hospitals* 60(13):31-2, July 5, 1986.

Fisch, C.: Electrocardiography and vectorcardiography. In Braunwald, E., *Heart disease*. 3rd ed. Philadelphia, PA. W.B. Saunders, 1988.

Garrison, L.P.: Cost containment and incentives for technology. *Health Affairs* 5(2):46-58, 1986.

Goodwin, J.C., and Dolan, E.G.: *Economics of public policy*. 3rd ed. St. Paul, MN. West Publishing, 1985.

Health Care Financing Administration, Health Standards and Quality Bureau. Monthly peer review organization data summary. Baltimore, MD., 1986.

Hobler, K.E., et al.: Excessive length of stay reduction without adverse effect on quality. *Quality Review Bulletin*, 11:239-241, 1985.

Hsia, D.C., Krushat, W.M., Fagan, A.B., et al.: Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. *New England Journal of Medicine* 318:352-355, 1988.

Kahkonen, D.M.: DRGs impact on medical care. *Michigan Medicine* 84:403, 1985.

Leamer, E.E.: Sensitivity analysis would help. *American Economic Review* 75:293-307, 1985.

Lee, T.H., and Goldman, L.: Serum enzyme assays in the diagnosis of acute myocardial infarction. *Annals of Internal Medicine* 105:221-233, 1986.

Mantel, N., and Haenszel, W.: Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* 22:719-748, 1959.

Mason, G.R.: Evaluating existing rate structures. *Top Health Care Financing* 14:36-49, 1987.

Matek, W., Guggenmoos-Holzmann, I., and Deming, L.: Follow-up of patients with colorectal adenomas. *Endoscopy* 17:175-181, 1985.

Mills, D.H.: *Report on medical insurance feasibility*. California Medical Association. Sacramento, CA. 1977.

Munoz, E., Goldstein, J., Lory, M.H., et al.: DRG hospital payment system, surgical readmissions, and cost containment. *American Surgery* 56:683-687, 1990.

Newhouse, J.P.: Do unprofitable patients face access problems? *Health Care Financing Review* 11(2):33-42. HCFA Pub. No. 03292. Office of Research and Demonstrations. Health Care Financing Administration. Washington. U.S. Government Printing Office, Winter 1989.

Piltch, C.: *Utilization and quality control peer review organization (pro) program—Program effectiveness*. Control No. OAI-88-00572. Office of Inspector General, U.S. Department of Health and Human Services. Boston, MA., 1988.

Poister, T.H.: *Public program analysis—Applied research methods*. Baltimore, MD. University Park Press, 1978.

Price, K.F.: Pricing Medicare's diagnosis-related groups—Charges versus estimated costs. *Health Care Financing Review* 11(1):79-90. HCFA Pub. No. 03286. Office of Research and Demonstrations. Health Care Financing Administration. Washington. U.S. Government Printing Office, Fall 1989.

Prospective Payment Assessment Commission: *Medicare prospective payment and the American health care system: Report to the Congress*. Washington, D.C. June 1990.

Public Health Service and Health Care Financing Administration: *International Classification of Diseases, 9th Revision, Clinical Modification*. DHHS Pub. No. 80-1260, Public Health Service. Washington. U.S. Government Printing Office, Sept. 1980.

Samuelson, P.A., and Nordhaus, P.A.: *Economics*. 12th ed. New York, NY. McGraw-Hill, 1985.

Siu, A.L., Sonneberg, F.A., Manning, W.G., et al.: Inappropriate use of hospitals in a randomized trial of health insurance plans. *New England Journal of Medicine* 315(19):1259-1266, 1986.

Stern, R.S. and Epstein, A.M.: Institutional responses to the prospective payment system base on diagnosis-related groups. *New England Journal of Medicine* 312(10):621-627, 1985.

Stokey, E., and Zeckhauser, R.: *A primer of public policy analysis*. New York. Norton, 1978.

Thompson, P., and Krushat, W.M.: *Report to Congress—Financial arrangements between physicians and health care businesses*. OEI-12-88-01410. Office of Inspector General, U.S. Department of Health and Human Services. Washington, D.C., 1989.

United States Code: U.S.C., 42, subsection 1320a-7(b)(6)(B) of the Social Security Act, 1989.

U.S. House of Representatives, Select Committee on Aging: *Sustaining quality health care under cost containment*. Committee Publication 99-499. Washington. U.S. Government Printing Office, Feb. 16, 1985.

U.S. House of Representatives, Select Committee on Aging: *Out "sooner and sicker"—myth or Medicare crisis*. Committee Publication 99-591. Washington. U.S. Government Printing Office, Apr. 10, 1986a.

U.S. Senate, Finance Committee: *Examination of quality of care under Medicare's prospective payment system*. Washington. U.S. Government Printing Office, June 23, 1986a.

U.S. House of Representatives, Ways and Means Committee: *Medicare Quality Protection Act of 1986.* Washington. U.S. Government Printing Office, Apr. 23, 1986b.

U.S. Senate, Special Committee on Aging: *Effects of the prospective payment system on quality of care for Medicare patients.* Washington. U.S. Government Printing Office, Jan. 7, 1986b.

U.S. Senate, Special Committee on Aging: *Quality of care under Medicare's prospective payment system.* Washington. U.S. Government Printing Office, Sept. 26, 1985.