A RAND NOTE

Causes and Effects of Expensive Hospitalizations

Emmett B. Keeler

RAND

The research described in this report was sponsored by the Health Care Financing Administration, U.S. Department of Health and Human Services under Cooperative Agreement No. C-98409/9-04.

Published 1991 by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

A RAND NOTE

N-3205-HCFA

Causes and Effects of Expensive Hospitalizations

Emmett B. Keeler

Prepared for the
Health Care Financing Administration,
U.S. Department of Health and Human Services

RAND

PREFACE

This Note was prepared as part of the RAND/UCLA/Harvard Health Care Financing Administration's research on reimbursement policy under Cooperative Agreement C-98409/9–04. It should interest health economists and policymakers concerned with the effects of outlier payments on hospital behavior, on the causes of unusually expensive cases, and on the distribution of justifiable but uncompensated cost differences between hospitals. Related work includes *A Description of Expensive and Long-Staying Patients,* by Sally Trude and Grace Carter (The RAND Corporation, N-2998-HCFA, October 1989) and *Insurance Aspects of DRG Outlier Payments,* by Emmett Keeler, Grace Carter, and Sally Trude (The RAND Corporation, N-2762-HHS, 1988). An editorial based on this work, entitled "What Proportion of Hospital Cost Differences Is Justifiable?" by Emmett B. Keeler, will appear in a forthcoming issue of *The Journal of Health Economics.*

SUMMARY

Medicare's Prospective Payment System (PPS) pays hospitals a fixed amount per case to give them incentives to economize. This system could give rise to problems of access and underservice for sick (i.e., expensive-to-treat) patients and losses to hospitals that treat high proportions of such patients.

The actual costs of treatment depend on patient needs, treatment intensity, and the hospital's efficiency in providing care and services. Ideally, hospital payments should reflect the costs for an efficient hospital to supply patient needs at an adequate standard of care. Adjustments to payments are already made for some factors that influence costs, including, especially, patients' diagnoses. Still, differences between average hospital payments and costs remain, and the distribution and causes of these differences (inefficiency, higher quality and intensity of services, or sicker patients) are of great interest. Indeed, health economists have argued for some cost sharing on the part of payers for cost differences resulting from unaccounted patient sickness and other "justifiable" factors. Unfortunately, cost sharing may subsidize inefficiency and services of questionable value: With reduced pressure to economize, hospitals may remain somewhat inefficient in providing care.

To address these issues, we analyzed previously collected data on costs, sickness, process, and outcomes from a large, nationally representative sample of hospitalized Medicare patients. We originally collected these data to study the PPS's effects on quality of care. We were able to match 87 percent of the 7156 post-PPS medical records with one of five diseases to Health Care Financing Administration (HCFA) files to get information on charges, outlier payments, assigned diagnosis-related group (DRG), and subsequent mortality.

To study the relation of sickness to cost/payment differences, we transformed charges to costs and we defined standardized (for payment factors) costs as costs standardized by dividing the payment adjustments for DRG weights, a wage index, disproportionate share, and teaching status. (We present a regression analysis showing that the current payment adjustments match costs fairly well.) For each disease, we had collected 60–80 disease-specific variables to measure the patients' sickness at admission. Previously these variables had been used to develop a scale that predicted mortality

within 30 days of admission. Here, we used regression methods to find scales that predicted standardized costs for each disease. We developed linear cost-sickness scales that included the previous death-sickness scale and 5–12 other variables, depending on disease. These scales predicted 10–15 percent of the variance in the log of standardized costs at the individual level, depending on disease.

Improved process of care was associated with higher costs, showing that the costs of providing higher quality were not recaptured by the cost savings of less complications of bad care. Quick death reduced costs, but nonfatal in-hospital complications greatly increased costs. With all these variables in the regression, 30–40 percent of the variance in the log of individual standardized costs was explainable, depending on disease.

We also looked at the determinants of average hospital costs/payments. Average sickness and process were important determinants, but only an indicator for urban government hospitals (which had considerably lower costs) among many hospital-level variables we tried (e.g., location, teaching status, size) was important.

Under certain assumptions, the optimal payment formula uses a weighted average of national average costs and hospital-specific costs, with the weight on hospital-specific costs set to be the R-squared of the regression of average hospital costs on "justifiable" factors (Pope, forthcoming). If only sickness at admission is justifiable, the weight is 15 percent. If quick death and nonfatal complications are included, the weight rises to 29 percent. Because our study did not measure many justifiable variables that affect costs, these are lower bounds.

Now that the switch to prospective payment is complete, outlier payments are the only cost-sharing program for acute-care hospitals. Outlier payments are given to hospitals for extremely expensive stays; in effect, they insure the hospitals against a portion of their risk. They have the form of a very large deductible—HCFA pays much of the costs that exceed a limit that depends on the diagnosis. The determinants of outlier cases and the effects of outlier payments exemplify the issues of payer cost sharing with hospitals in extreme form. We had hoped to analyze their determinants and effects to shed light on the desirability of such cost sharing.

Unfortunately, our original plan to estimate the effects of outlier payments on hospital behavior was thwarted by the data. Two problems derived from our data coming from the 1985–1986 period when outlier payments were being phased in. First, outlier incentives were weak at that time—outlier payments covered only 30 percent of the costs of care over the threshold; second, many cases that should have been outliers were not

coded as such in the Bill Retrieval File. As a result, although we could not find effects of outlier payments on hospital behavior at that time, we cannot infer that outlier incentives do not influence behavior today. The problems in the transitional data may have prevented us from discovering a true effect.

To study the determinants of outliers, we defined synthetic outliers: cases with cost/payment ratios greater than three times the average, or cases for whom hospitals actually received outlier payments. Synthetic outliers made up 4.5 percent of the cases in our five diseases. The patient and hospital determinants of outliers are very similar to those of expensive cases generally, with the exception of urban government hospitals and in-hospital deaths which, despite their association with lower costs, generally are not underrepresented among outliers.

ACKNOWLEDGMENTS

We would like to thank Grace Carter and Judith Lave for many helpful suggestions, Katherine Kahn for the data from the original Prospective Payment System study, Stan Bentow for programming support, and Martha Cooper for typing.

CONTENTS

TABLES

I. INTRODUCTION

The institution of Medicare's Prospective Payment System (PPS) based on diagnosis-related groups (DRGs) was accompanied by dramatic decreases in lengths of stay (Guterman et al., 1988). These decreases testified to the importance of the new incentives to economize on the costs of hospital stays. However, the system of fixed payments per case potentially could give rise to new problems of access and underservice for very sick (and hence very expensive-to-treat) patients and to losses and risks to hospitals that treat unusually high proportions of such patients.

The actual costs of treatment depend on patient needs, treatment intensity, and the hospital's efficiency in providing care and services. Ideally, hospital payments should reflect an efficient hospital's costs to supply patient needs at an adequate standard of care. Adjustments to payments are already made for some factors that influence costs, including the local wage level, teaching status, a disproportionate share of poor patients, and, most important, patients' diagnoses. After these adjustments have been made, adjusted payments are not exactly equal to costs, and hospitals will have different average (costs-payments). The distribution and causes of these differences—i.e., the proportion of higher costs resulting from inefficiency, from higher quality and intensity of services, or from sicker (in ways not captured by DRGs) patients—is of great interest for policy.[1]

Health economists have argued for some cost sharing on the part of payers for differences resulting from patient sickness at admission. Such arguments can be made on both efficiency (Ellis and McGuire, 1986) and equity grounds (Pope, forthcoming; Goodall, forthcoming). Unfortunately, cost sharing may subsidize inefficiency and services of questionable value: With reduced pressure to economize, hospitals may remain somewhat inefficient in providing care.

Now that the switch to prospective payment is complete, outlier payments are the only cost-sharing program for acute-care hospitals. Hospitals receive outlier payments for extremely expensive stays, in effect insuring them against a portion of their risk (Ellis and McGuire, 1988). In principle, one could look at outliers to study both the

[1]Here we consider only briefly other reasons for differences between costs and payments, such as imprecise adjustment for wages or the lack of adjustments for other input-factor costs.

determinants of hospital expense (in extreme form) and the effects of the Health Care Financing Administration's (HCFA's) cost sharing on hospital behavior. We present some analyses of outliers here, but because of problems with the coding of outliers in the data, the rareness of outliers, and interest in the costs of hospital care generally, we have extended our analyses of the costs of care to all cases.

In this Note, we present analyses of the determinants of costs. We take advantage of clinically detailed data gathered previously from the medical records of a nationally representative sample of over 14,000 patients with five diseases. We matched the post-PPS half of these data to administrative HCFA data on payments. We address two main sets of issues. First, what additional sickness variables (besides the DRG) predict costs of treatment, how much of the variance in costs do these variables predict, and what proportion of differences in costs between hospitals are caused by sickness (and are hence unavoidable by hospitals that take expensive patients)? Second, what else besides sickness causes higher average costs? In particular, does higher quality necessitate higher costs, or does it pay for itself through fewer complications (Haley et al., 1987)? Do the current hospital payment adjusters for location, teaching status, and disproportionate share match costs well, and are there other good hospital-level predictors for costs? Are the hospital-level payment adjusters simply proxies for sickness and quality, or do they represent other independent effects?

In Sec. II, we describe the data and methods. Section III contains a description of patient characteristics that determine expensive cases. Earlier approaches have used bigger data sets with limited information (age, diagnosis, hospital characteristics) (Trude and Carter, 1989) or have looked retrospectively at a few hospitals (Schroeder et al., 1979, 1981; Zook and Moore, 1980). Here, we use the detailed information on sickness at admission to derive scales that predict expenses and allow us to discriminate between the effects of patient severity, mishaps, and hospital characteristics on the average costs of care.

Section IV contains the hospital-level analyses of costs. We look at the relation of costs to the factors that are currently used to adjust payments and to other hospital characteristics. Recently, some health economists have argued that prospective payments should be based on a blend of hospital-specific and national average costs, with the weights depending on the proportion of costs resulting from "justifiable" factors, such as patient severity (Pope, forthcoming; Goodall, forthcoming). We present their arguments and interpret our results in this framework.

Section V contains an analysis of the determinants of outliers. Because many cases that should have been outliers were not so coded in our data, we defined and analyzed synthetic outliers: cases with cost/payment ratios greater than three, or cases for whom hospitals actually received outlier payments. Synthetic outliers made up 4.5 percent of the cases in our five diseases. We show that the patient and hospital determinants of these outliers are very similar to those of more expensive cases generally.

Appendix B contains an analysis of the relationship between outlier status and quality of care. The analysis tries to see if outlier payments have important effects on the care of hospitalized patients. Unfortunately, in 1985–1986, the time selected for the post-PPS half of the sample, outlier payments were small, and both hospitals and fiscal intermediaries were still confused about which cases should be paid. Both factors dilute the potential impact of outlier payments on incentives and render less interesting the finding that cost sharing has no effect on hospital behavior.

II. METHODS

STUDY DESIGN AND SAMPLE

The original study attempted to test whether quality of care for hospitalized Medicare patients fell as a result of the introduction of PPS. It compared patients who were hospitalized before the implementation of PPS (calendar years 1981 and 1982) with patients hospitalized after PPS was introduced (July 1, 1985, to June 30, 1986) (Draper et al., forthcoming). In the present study, we will only use patients from the post-PPS period. The study included only Medicare patients age 65 years and older.

We studied patients who were hospitalized in one of 296 acute-care general hospitals from 30 cities or towns in five states, each from a different region of the United States. We selected the hospitals to represent the national patient cohort with respect to city size, percentage of Medicare patients, hospital size, teaching intensity, and type of ownership. We oversampled hospitals caring for poor patients to address questions about care for the poor. Estimates weighted to reflect our oversampling differ little from unweighted estimates in our major analyses; we present only unweighted results here.

Patient Sample

We start with data from a sample of 7156 patients hospitalized in 1985 and 1986 with one of five diseases: congestive heart failure (CHF), acute myocardial infarction (AMI), pneumonia (PNE), cerebrovascular accident (CVA, or stroke), and hip fracture (HIP). Patients with these diseases represented 19 percent of Medicare discharges and 32 percent of Medicare deaths within 30 days of admission in FY 1986. We initially selected patients for study if they had an ICD-9-CM code consistent with one of the study diseases. If such a code was present, we screened the patient's medical record to verify that the study disease was the reason for hospital admission. Only patients at least 65 years of age who were admitted with the study's diseases were eligible for inclusion. We present details of the study's sampling plan elsewhere (Draper et al., forthcoming).

Mortality Data

We used the medical record as our source of in-hospital mortality information and existing HCFA files to determine mortality status after the patient's discharge. By using the patient's last name, first name, date of birth, and health insurance claim number from

the medical record, we were able to accurately match 93 percent of the post-PPS patients in our study sample to the HCFA Health Insurance Master File (HIMF).

We tried to ascertain whether the 7 percent of patients we could not match were different in important ways from the other 92 percent of patients. The unmatched patients were sicker, on average, than were matched patients, and 19 percent had died in-hospital.

Charge Data

To study charges we used the patients we matched to the HIMF and matched them to HCFA's Bill Retrieval File of hospitalizations. Of the 93 percent of post-PPS patients for whom we had accurate mortality data, we were able to accurately match 94 percent in the Bill Retrieval File, yielding an overall matching success rate of 87 percent (see Table 2.1). The resulting file allowed us to measure the charges and outlier payments and to find the assigned DRG for the admissions we had abstracted.

The distribution of total charges was very skewed, so for the individual-level analyses we used the log of charges, dropping the six cases with no charges. We transformed charges to costs by adding the log of the 1984 cost-to-charge ratio. In the 13 hospitals for which the cost-to-charge ratio was missing, we used 0.66, the 1984 national average. (Because the log of the cost-to-charge ratio has a coefficient of about −1 in the regression of log charges on it and other explanatory variables, the results on predicting log charges would be very similar to those on log costs presented below.)

Controlling Costs for Factors Currently Used to Adjust Payment

Current payment regulations reflect the fact that a hospital's costs depend on the wages it must pay, the diseases of the patients it treats, the amount of teaching its staff members undertake, and whether it treats a "disproportionate share" of poor people.

Table 2.1

SAMPLE FOR COST ANALYSIS

Sample Category	CHF	AMI	PNE	CVA	HIP	Total
Medical records	1465	1437	1408	1442	1404	7156
Match to HIMF	1367	1329	1321	1342	1295	6665
Match to nonzero charge data[a]	1279	1237	1227	1250	1228	6221

[a]Six cases with zero charges in the HCFA files were dropped.

These four factors are used to adjust payments. Because we wanted to study other determinants of costs not currently used by HCFA to adjust payments, we standardized costs to control for these cost factors that were already included. We defined the standardized (for payment factors) costs per case to be the costs divided by the payment factors: the FY 1987 wage index, the 1986 DRG weights, and the disproportionate-share and teaching adjustments. Taking logs, we obtain

$$\log(\text{standardized costs}) = \log(\text{costs}) - \log(\text{DRG weights}) - \log(0.2085 + 0.7915 * \text{1987 wage index}) - \log(\text{dsh} + 1.5 * [(1 + \text{irb})^{0.405} - 1],$$

where dsh is the 1987 disproportionate-share factor, and irb is the 1984 interns/residents-to-beds ratio.

Standardized costs are closely related to the hospital's loss on the case.[1] The log (standardized costs) is almost a linear function of the residuals of log (costs) when regressed on the four payment variables (see Table 4.1, below, for coefficients). The average correlation across the five diseases of log (standardized costs) and the residuals of a regression of log (costs) on payment factors was 0.995. Thus, when looking at other determinants of costs besides the current payment adjustors, whether we regress log (costs) on the payment variables and other variables or log (standardized costs) on the other variables by themselves doesn't matter.

Assuming that the current adjustments are correct (an assumption we test in Sec. IV), the standardized costs are the costs standardized to a hospital facing average local wages, a DRG average case mix of one, no teaching, and no disproportionate share. Because the five diseases in the sample are serious, the average DRG weight was 1.5, and the standardized costs have a mean of $3130, as opposed to a mean of $4600 for unstandardized costs.

[1]Indeed, because payments = standardized payment amount * payment factor, the standardized costs = (cost/payment) * standardized payment amount. This in turn is proportional to 1/(1 + margin). Cases or hospitals with large standardized costs are losers; cases or hospitals with small standardized costs are winners.

SICKNESS AND OUTCOME VARIABLES

Measuring Sickness at Admission

To take into account patients' sickness at admission, we collected 60–80 disease-specific variables from the medical record that described the patient's acute and chronic morbid and comorbid conditions (Kahn et al., 1988a; Kosecoff et al., 1988; Roth et al., 1988; Rubenstein et al., 1988; Sherwood et al., 1988). Table 2.2 lists variables collected for all five study diseases, as well as variables collected specifically for each disease.

In developing measures of sickness at admission, we paid special attention to previously developed severity measures, including those developed to predict death for intensive-care unit (ICU) patients (Knaus et al., 1985) and measures for other particular groups of patients with long and useful clinical histories (Killip and Kimball, 1967; Norris et al., 1969; Goldman et al., 1977).

We used literature review, clinical judgment, and disease-specific consensus panels to identify a list of variables that have been considered important clinical predictors of death, nonfatal complications, readmissions, or prolonged nursing-home residence. We included variables if they accurately described the patient's status at the time of hospital admission, were frequent enough to be worth collecting, and were reliably recorded in the medical record (Kahn et al., 1988a).

We used disease-specific abstraction forms to guide the collection of data about sickness at admission from the medical records of hospitalized sampled patients. We included measures of patient sickness at admission based on the reason for the hospital admission (morbid conditions) and on other conditions that might influence outcomes (comorbid conditions). For each patient, we collected data about acute and chronic morbid and comorbid diseases, function, the number of body systems with pathology, and the APACHE II Acute Physiological Score (APS) variables (Knaus et al., 1985). For patients hospitalized with AMI, we also collected data that can be used to score the Norris Coronary Prognostic Index and the Killip scale (Norris et al., 1969; Killip et al., 1967). For hip fracture patients, we collected data for the Goldman Preoperative Risk Index (Goldman et al., 1977).

Data were collected by 52 data collectors (nurses or medical-records technicians) with prior experience in medical-record abstraction. Data availability and interrater reliability are described elsewhere (Draper et al., forthcoming). Abstractors collected data about sickness at admission from physician and nurse notes (including emergency

Table 2.2

COMPLEXITY VARIABLES STUDIED TO MEASURE SICKNESS AT ADMISSION

CHF	AMI	HIP	PNE	CVA	All Diseases[a]
Acute Conditions					
Recent angina (past week)	Shock[b] Location of infarction[c] Extent of infarction (e.g., transmural)[c] Difficulty using limbs	Fracture location by M.D. report (e.g., intertrochanteric) Fracture stability by M.D. report (e.g., garden stage, displacement) Atrial arrhythmias[d] Ventricular arrhythmias[d] Heart block by EKG[d]	Recent influenza Respiratory distress Wheezing Empyema Lung abscess Meningitis Pericarditis Endocarditis	Hemorrhagic infarction Recent angina (within 1 week) Recent CHF (within 1 week)	Shortness of breath Chest pain Arrest Intubation (except HIP)
Demographics					
					Age,[c,e] sex, race
Prior Chronic Conditions					
Noncompliance with therapy Difficulty using limbs		Status post pneumonectomy Deep venous thrombosis Pulmonary embolism Parkinson's disease Paget's disease Osteoarthritis Rheumatoid arthritis Osteoporosis Prior contralateral hip fracture Hip malignancy (new diagnosis this admission)	Immunocompromise Multiple myeloma Status post splenectomy Respiratory failure Pneumonectomy Aspiration Dysphagia Thoracic neuromuscular disease Recent thoracic or abdom. surg. Interstitial lung disease Nasogastric tube before admission Tracheostomy Lung cancer (new diagnosis during admission) Morbid obesity Current alcoholism Prior hospitalization for lung disease	Immunocompromise Deep venous thrombosis Malnutrition Nasogastric tube before admission Current alcoholism Morbid obesity Temporal arteritis Systemic lupus erythematosis Status post carotid endarterectomy	Valvular disease Angioplasty or coronary bypass surgery Angina[c] Pacemaker Atypical chest pain (except HIP and PNE) Myocardial infarction[c] CHF Hypertension (except HIP and PNE) Arrhythmia (except CHF and AMI) Arterial disease (except HIP and PNE) Emphysema, COPD Use of home oxygen (except CVA) Cancer diagnosis Cirrhosis[e] Diabetes Renal failure Cerebrovascular accident Psychosis or depression (except PNE) Malnutrition (except CHF and AMI) Current smoking (except HIP) Hospitalization within 6 months (except AMI)

Table 2.2 (continued)

CHF	AMI	HIP	PNE	CVA	All Diseases
			Function		
					Dementia Urinary continence Ambulatory status From nursing home
			Radiographic		
		Fracture location (e.g., intertrochanteric) Fracture displacement Fracture comminution Fracture impaction Fracture stable Fracture valgus Fracture varus Stress fracture		Increased intracranial pressure by CAT scan Lacunar infarc. by CAT scan Hemorrhagic infarc. by CAT scan	CHF by chest x-ray[b,c] PNE by chest x-ray (except AMI)
			Physical Examination		
	Lung not clear to auscultation (e.g., rales) CPK isoenzymes LDH isoenzymes Cardiomegaly[c]			Speech deficit Facial deficit Visual deficit Motor deficit Sensory limb deficit Unresponsive to painful stimuli[g] Unresponsive to touch[g] Unable to follow commands[g] Decorticate posturing[g]	Coma (except HIP)[f,g] Confusion or altered neurologic status[d,f,g]
			Laboratory Studies		
Digitalis level Positive blood culture			Positive blood culture Positive urine culture	Sedimentation rate	Oxygenation (A-aD02 or p02)[f] Arterial pH[f] Serum sodium[f] Serum potassium[f] Serum creatinine[f] Blood urea nitrogen[f] Venous bicarbonate (except HIP)[f] Hematocrit or hemoglobin[f] White blood count[f] Serum CPK (except HIP) Serum SGOT Serum bilirubin Serum albumin Missing laboratory values

Table 2.2 (continued)

CHF	AMI	HIP	PNE	CVA	All Diseases
			Prior Medications		
Beta blocker			Antibiotics	Anticoagulants	Digitalis (except CVA)
Other left vent. depressant			Aminophylline	Antiarrhythmic	Insulin
Antiarrhythmic					Systemic steroids (except AMI)
			Vital Signs		
					Temperature[f]
					Systolic blood pressure[c,f]
					Diastolic blood pressure[f]
					Heart rate[f]
					Respiratory rate[f]

[a]We measured the body system count for each disease by counting the number of each of the following body systems evincing pathology: cardiac, pulmonary, hepatobiliary, endocrine, arterial, renal, neurologic, psychiatric, oncologic, immunologic, hematologic, musculoskeletal, and nutrition.

[b]These variables are part of the Norris Coronary Prognostic Index.

[c]These variables are part of the Killip scale.

[d]These variables are part of the Goldman Preoperative Risk Index.

[e]These variables are part of the APACHE II Chronic Health Evaluation system.

[f]These variables are part of the APACHE II Acute Physiologic Score (APS) system. Systolic and diastolic blood pressure are combined in the APACHE APS into a mean blood pressure value ((2 × diastolic + systolic blood pressure)/3). The APACHE APS system uses venous bicarbonate only if no arterial blood gas value is available.

[g]These variables are part of the RAND CVA coma scale.

room notes), laboratory results, and radiographic reports. In all instances, we collected the first available data for the patient after admission. If the patient had missing laboratory or radiographic data on day 1, we accepted data from day 2 as a measure of sickness at admission because some patients are admitted late at night and day 2 values could represent their admission status.[2]

In addition to variables in the medical record, we looked at the possible sickness measures available from the administrative files. These included demographics, whether surgery was performed, coded comorbidities, and prior hospitalizations. Such variables are used in the HCFA mortality predictions and in Trude and Carter (1989) to predict costs. Despite the known problems in diagnosis coding (Jencks, Williams, and Kay, 1988), these variables have predictive power. We also created a file of these variables, but we reached our resource limits before we could perform a comparative analysis of administrative and medical-records data.

Methods for Developing the Cost-Sickness Scales for Each Disease

For analytic purposes, it is useful to combine the sickness indicators into scales that predict costs for each disease. Using such scales simplifies comparisons of the contribution of sickness to costs across diseases and hospitals, and the relative impacts of sickness, process, complications, and death. Creating such scales requires two steps: scaling the items and combining them. For both steps, we built on previous work to analyze mortality (Keeler et al., forthcoming).

By scaling the items, we mean assigning numerical values to the various possible results for a particular item, e.g., for low systolic blood pressure, deciding how many points a blood pressure of 70 should get, given that normal blood pressure (140–160) receives no points and a blood pressure of 120 receives one. In addition, one must decide what score to assign to missing values. In this work, we simply used the item scalings we had developed previously. Thus, for all variables except age, blood urea nitrogen (BUN), and temperature, normal values get a score of zero, and the maximum score assigned for the sickest people is also the range of the score. For variables defining the patients' medical and functional history, we assumed that patients did not have a condition unless it was mentioned in their medical records.

[2]We did not include the presence of a "do not resuscitate" (DNR) order as a measure of sickness. Although related to sickness, it also reflects a (process) decision of the provider that could confound a pure sickness scale.

The second step in creating a scale is selecting the items to combine in the scale, together with their weights. We had previously developed sickness scales for each disease to predict 30-day mortality and 180-day mortality, called sev30 and sev180.[3] For each disease, sev180 was the sum of sev30 and a subscale, add180, which contained many chronic-condition indicators. We used sev30 and add180 here together with other sickness items to create the new cost-sickness scale for each disease.

We used a preliminary screen to reduce the number of sickness-at-admission variables that would be used to predict costs. The dependent variable was always the log of costs. For each disease, we regressed the log of costs on the logs of the four payment adjustments and the sickness/mortality scales, sev30, and add180, together with subsets of the 60–80 other sickness-at-admission variables. The other sickness-at-admission variables were kept for the final regressions that defined sickness according to a rule that combined judgment with statistical evidence. We kept variables for possible inclusion in the scale if they were either strong predictors ($t > 2.5$) or medium predictors ($t > 1.5$) and we expected a strong connection with cost.

Finally, for each disease we used backward stepwise regression of log costs on the payment adjustors and all the candidate variables (including sev30, add180, and sev30 squared) to weight variables. We dropped insignificant variables.[4] The final weights of the sickness variables were the regression coefficients rounded to three decimal places. The rounding of the sickness weights in the scales had no effect on R-squared (at least up to four decimal places). To estimate conservatively the ability of these cost-sickness variables to predict costs,[5] we divided the sample into ten equal parts. The reported R-squared value is the average of the squares of the simple correlation between outcomes in each tenth and predictions based on the other 90 percent of cases.

In-Hospital Adverse Outcomes and Length of Stay

The in-hospital adverse outcomes come in three parts: major complications, general sickness at day 3 (a scale of seven vital signs and symptoms), and new lab abnormalities. We did not include adverse outcomes that occurred on the day of death

[3]Although the name is the same, sev30 for, say, hip fracture patients is different from sev30 for pneumonia or any other disease. Because analyses are done disease by disease, the common names are not a problem.

[4]We dropped two marginally significant variables ($t < 2.5$) with improbable negative associations (nasogastral tube for PNE, low systolic blood pressure for CVA).

[5]Because the cost-sickness scales are developed on the same data used to estimate their predictive ability, the simple R-squared is an overestimate of how well they would do on other data.

but did include adverse outcomes that were followed by later in-hospital death. Thus, it would be more precise to call them "not immediately fatal" adverse outcomes. Here, we use the overall nonfatal complications scales that were derived for each disease to supplement in-hospital death as a measure of bad outcomes. They were derived by regressing mortality within 180 days of admission on the major complications, the previously scaled signs and symptoms, and previously scaled new lab abnormalities, adjusting for sickness at admission.[6] They are scaled here so that an increase of 0.01 in the complications scale is associated with a 1 percentage point increase in 180-day postadmission death, after adjusting for sickness at admission.

We also collected length-of-stay data from the medical records. This variable is somewhat endogenous (under the control of the hospital), but there is some interest in seeing whether factors that influence cost do so through cost per day or length of stay.

Measuring Instability at Discharge

Our measure of instability at discharge was based on data from the day of discharge or (since most patients are discharged in the morning) the day before discharge. Instability variables are designed to identify patient problems present at discharge that (1) clinicians generally agree should be corrected before discharge and (2) are likely to cause poor outcomes if they are not corrected. They reflect correctable problems that occurred in the stay and were not present at admission. In the full PPS sample, 16.5 percent of patients had at least one of the 8–10 new problems per disease that were counted as instability (Kosecoff et al., forthcoming).

Measuring Preadmission Residence and Discharge Destination

We used the patient's medical record as the source of both preadmission residence and discharge destination, labeling origin and/or destination "home" if so indicated in the physicians', nurses', or therapists' notes or in administrative data. For both origin and destination we defined "nursing home" as a skilled nursing facility, intermediate-care facility, extended-care facility, or a nursing home without other description. We listed retirement homes, sheltered housing, congregate housing, halfway housing, and board-and-care facilities as "residential homes."

[6]In the original study, data on in-hospital complications were collected to be used to measure serious long-term health effects of treatment. Thus, we did not collect data on complications, such as nosocomial infections, that might have strong effects on costs but not on mortality. Because the complications scale was tailored to predict death, the results presented below underestimate the effects of complications on costs.

Measuring Process by Explicit Criteria

To create clinically valid process measures, we performed a literature review and consulted experts to identify process measures that made a difference in patient outcomes. These measures were then reviewed by disease-specific panels of physicians from both community and university hospitals to decide whether the criteria made clinical sense and would not be vulnerable to variations in recording by year, state, or hospital type. When the panel believed there were several good strategies for diagnosing or treating a patient, the process criteria assigned credit if any of the strategies were followed.

We developed and pilot-tested disease-specific abstraction forms to collect the medical-records data. In addition to studying process measures applicable to all diseases, we also gathered disease-specific process variables. For each disease, we collected process data on the first two hospital days; the first time a clinical problem arose, regardless of the hospital day; and the day before and day of discharge. For the four medical diseases, we also studied process of care on hospital day 3. For patients with hip fracture, we collected data on the days surrounding surgery. The instruments for measuring process have been published elsewhere (Kahn et al., 1988b; Kosecoff et al., 1988; Roth et al., 1988; Rubenstein et al., 1988; Sherwood et al., 1988).

Using clinical judgment, we grouped process criteria and then tested our groupings by comparing them to those suggested by a Likert scaling model. These methods produced five conceptual process scales: physician cognitive, nurse cognitive, technical diagnostic, technical therapeutic, and monitoring with ICU or telemetry. We combined these five scales to create the overall process scale analyzed here.

To produce the component scales, we had to combine some process measures that were applicable to all patients with those that were applicable to subsets of patients. To avoid a bias when combining criteria to form scales, we standardized all process criteria to have mean of zero. Using this method meant that each patient who received average process would have an expected process score of zero, regardless of the number of process-score items applicable to the patient.

Elsewhere, we have given details of the measurement of quality of care using explicit process and have shown that higher in-hospital process scores were associated with lower postadmission mortality, after adjusting for sickness at admission (Kahn, Rogers, and Rubenstein, forthcoming).

III. WHAT PREDICTS EXPENSIVE CASES?

In this section, we analyze the determinants of costs of individual hospital stays. We discuss hospital-level factors in Sec. IV.

THE COST/SICKNESS-AT-ADMISSION SCALE

The sickness scales to predict costs are shown in Tables 3.1–3.5. The first two columns of these tables show the coefficients and t-values for the sickness variables from a regression of log (standardized costs) on these variables together with the payment variables (whose coefficients are not shown). The last two columns of the tables give additional information on the variables in the scale. The scales are simply the weighted sum of the variables shown, with the weights given by the regression coefficients. Because the scales are linear predictors of the log of standardized costs, a 0.01 increase in the cost-sickness scale is associated with a 1 percent increase in standardized costs, on average.

For the diseases with high in-hospital mortality (AMI, CVA, and pneumonia) the sickness-at-admission scale that predicts 30-day mortality enters quadratically, because many of the sickest people die quickly and are therefore inexpensive. The maximum cost for these diseases is at the sickness-at-admission level corresponding to a 10 percent

Table 3.1

COST-SICKNESS SCALE FOR CONGESTIVE HEART FAILURE

Variable	Coef.	t	Mean	Maximum Value
Sickness-at-admission scale	0.0089	3.64	32.45	NA
Blood oxygen score	0.061	4.75	0.91	4
Acute complexity sum[a]	0.056	5.45	2.43	15
Prior COPD	0.127	2.77	0.24	1
Albumin score	0.068	2.11	0.16	3
Body-system count	0.036	2.30	2.94	8

[a]Sum of X-ray CHF score, X-ray pneumonia score, CPK score, digitalis score, SGOT score, bilirubin score, arrest in emergency room, intubation day 1, blood-culture fever score.

Table 3.2

COST-SICKNESS SCALE FOR ACUTE MYOCARDIAL INFARCTION

Variable	Coef.	t	Mean	Maximum Value
Sickness-at-admission scale	0.008	1.80	25.82	NA
Sickness squared/100	−0.030	−4.96	8.56	NA
Blood-gas oxygen score	0.060	3.97	0.58	4
Serum creatinine score	0.050	2.99	0.76	4
Hematocrit	0.061	2.02	0.25	4
APACHE chronic HE score	0.059	2.95	0.15	5
X-ray CHF score	0.035	2.11	1.45	4
SGOT score	0.151	3.66	0.24	1
Psychotic	−0.162	−3.23	0.06	2
Demented	−0.168	−2.47	0.05	2
From nursing home	−0.045	−1.74	0.18	3
Body-system count	0.076	4.23	2.37	7

chance of dying within 30 days for AMI, 20 percent for pneumonia, and 22.5 percent for CVA.

In the five sickness scales, costs are generally more associated with acute problems (both those comprised by the sev30 scales, and the other acute problems shown

Table 3.3

COST-SICKNESS SCALE FOR PNEUMONIA

Variable	Coef.	t	Mean	Maximum Value
Sickness-at-admission scale	0.016	3.70	28.09	NA
Sickness squared/100	−0.023	−3.92	10.49	NA
Sickness added for 180-day death	0.017	4.31	6.75	NA
Pleural effusion score	0.18	3.09	0.13	1
Arrest in emergency room	0.12	2.71	0.06	3
Blood-gas oxygen score	0.08	6.26	1.35	4
Potassium abnormality	0.10	2.68	0.23	4
CPK score	0.13	2.97	0.08	3
Positive blood culture	0.07	3.23	0.29	3
Positive urine culture	0.20	3.02	0.11	1
Any of 4 rare lung diseases	0.13	2.81	0.04	4
Prior COPD score	0.11	2.70	0.35	1
Current smoker	0.09	3.11	0.26	2
Alcoholism	0.13	2.18	0.06	2
From nursing home	−0.06	−2.47	0.56	2

Table 3.4

COST-SICKNESS SCALE FOR CEREBROVASCULAR ACCIDENT

Variable	Coef.	t	Mean	Maximum Value
Sickness-at-admission scale	0.039	7.05	19.35	NA
Sickness squared/100	−0.080	−8.31	5.41	NA
Sickness added for 180-day death	0.021	4.74	9.39	NA
Age	−0.0068	−2.24	78.01	113
Blood-gas pH score	0.081	2.24	0.13	4
Blood-gas oxygen score	0.071	2.97	0.28	4
Acute complexity score[a]	0.050	5.44	5.39	29
Erythrocyte sedimentation	0.167	3.46	0.06	3
From nursing home	−0.113	−4.45	0.34	3

[a]Sum of X-ray CHF score, X-ray pneumonia score, CPK score, digitalis score, SGOT score, bilirubin score, arrest in emergency room, intubation day 1, blood-culture fever score.

Table 3.5

COST-SICKNESS SCALE FOR HIP FRACTURE

Variable	Coef.	t	Mean	Maximum Value
Sickness-at-admission scale	0.015	6.13	7.419	NA
Comminuted fracture	0.132	4.77	0.337	1
Impacted fracture	0.083	2.51	−0.203	0
Fracture location scale	0.198	3.52	1.965	2
Ventric. arrhythmias EKG day 1	0.121	2.57	0.080	1
Osteoarthritis	0.133	2.85	0.082	1
Stress-fracture score	−0.941	−2.07	0.0008	1
From nursing home	−0.045	−3.78	0.616	3
Demented	−0.043	−2.94	0.552	3
Body system count	0.035	3.20	2.845	9

in the tables) than with chronic conditions. The acute complexity sum enters directly in CHF and CVA, and its components, such as arrest in the emergency room, pleural effusion, positive blood and urine cultures, and CPK scores are important for pneumonia. Laboratory and X-ray values indicating immediate problems add to costs for AMI. Hip fracture costs are related to variables associated with the difficulty of repairing the fracture.

For some diseases, a few selected chronic problems add to costs. Prior chronic obstructive pulmonary disease (COPD) and a low albumin score (a measure of malnutrition) were important for CHF. SGOT scores (a measure of liver problems) and the APACHE chronic health evaluation add to AMI costs. Prior COPD and smoking and rare lung diseases add to pneumonia costs. Erythrocyte sedimentation abnormalities were expensive for CVA.

The variable add180, which is added to sev30 to make the best predictor of 180-day postadmission mortality, appears in only two of the five diseases. The body system count that appears in three of the five diseases is a simple sum that counts the number of body systems with either chronic or acute problems. For all diseases but CHF, variables that point to a lower posthospital quality of life are associated with lower costs. These include whether the patient arrives from a nursing home (for four diseases) or is demented, psychotic, or alcoholic (variables that have lower costs in some diseases).[1]

For the medical conditions, a blood-gas oxygen abnormality is associated with higher costs. The blood-gas tests are often not done to patients who would show normal values: In fact, analysis done elsewhere shows that people who do not get these tests are less likely to die than people who do but have normal results (Keeler et al., forthcoming). Moreover, the tests are a marker for more technologically advanced care (they were done considerably more in 1985–1986 than in 1981, and are done considerably more in teaching than in nonteaching hospitals). To test whether abnormalities are a proxy for more high-tech, expensive hospitals or simply reflect an expensive problem case, we constructed variables giving the proportion of patients in each hospital missing blood-gas tests and missing blood-chemistry tests. These variables were never important predictors of costs, implying that patients with abnormal blood gases cost more to treat because they are sick, not because they have gone to a hospital that does many tests.

Process

We regressed the log of standardized costs on the six process subscales and the sickness scale. The physician and the nurse cognitive-diagnosis scales are based on very low-tech criteria, e.g., did they gather data (and record them in the medical record) about

[1] If the prognosis or posthospital quality of life is low enough, providers may add a DNR code to the medical record. Elsewhere, we have shown that DNR indicators are related to such variables. Although DNR orders are related to sickness, we did not allow an indicator for the presence of a DNR order to be part of the scale, because it is partly a decision (process) variable for the provider.

the patient's history, symptoms, and physical examination, including vital signs? These scales were the ones most strongly associated with better outcomes after adjusting for sickness at admission, probably because they apply to all patients, and most of the other process measures are of the "if patient is X, then do Y" type. The other process scales (proper use of the ICU, technical therapeutic, technical diagnostic, and checking of abnormal labs) might be expected to raise costs directly, even if they lowered them indirectly by making the patient less sick.

The relation of the components of process to costs for the five diseases is shown in Table 3.6. Because each process scale has standard deviation 1, each coefficient has a simple interpretation. For example, the 0.02 coefficient for physician cognitive process in CHF means that a patient for whom such process is 1 standard deviation better than average (at the 84th percentile) has costs 2 percent greater than if their process were average. Proper rechecking of labs and ICU/telemetry have the biggest effects on costs, which seems reasonable since fulfilling the criteria requires that the hospital spend resources for sick patients. It is more surprising that, except for hip fracture patients, the diagnostic components of process are also related to higher costs. The technical therapeutic process (measuring medications, physical therapy, etc.) is only significantly related to costs for AMI patients. (We did not collect such data for CVA patients.)

Some have speculated that higher-quality care could result in financial gains to the hospital (Haley et al., 1987), but these findings show that the gains do not come by saving money on the case at hand. Of course, a reputation for higher quality would be good for a hospital's morale and competitive position, everything else equal.

Table 3.6

REGRESSION OF LOG (STANDARDIZED COSTS) ON PROCESS VARIABLES

	CHF		AMI		PNE		CVA		HIP	
	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Physician cognitive	0.02	0.9	0.06	3.4	0.02	0.9	0.08	3.4	0.00	0.3
Nurse cognitive	0.04	2.1	−0.03	−1.7	0.06	2.9	0.06	3.1	0.01	0.6
Technical DX	0.00	0.0	0.12	6.1	0.06	3.1	0.05	2.4	0.04	3.1
Technical TX	0.03	1.4	0.06	2.8	0.02	0.9	NA	NA	0.01	0.9
ICU/telemetry	0.22	6.0	0.05	2.6	0.16	6.3	0.11	2.9	0.14	4.4
Labs checked	0.13	5.2	0.13	4.4	0.20	6.2	0.16	5.0	0.04	2.2
Sickness	0.92	12.1	0.78	9.7	0.71	10.7	0.87	12.5	0.93	11.8
N	1279		1237		1227		1250		1228	
R-squared	0.17		0.20		0.22		0.20		0.15	

In-Hospital Outcomes

Table 3.7 shows the effects of in-hospital outcomes on standardized costs. The outcomes include in-hospital death and our scale of adverse outcomes not immediately fatal (comprising major complications, symptoms and signs on day 3, and new bad laboratory values). In-hospital outcomes have a big impact on costs; the variance in log (adjusted costs) explained jumps by an average of 25 percentage points.

For all diseases, in-hospital death was associated with much lower costs. The average length of stay of those who died in-hospital compared to that for other patients was dramatically shorter for AMI (5.4 compared with 11.7 days) but not for the other diseases. The lower costs for death are a partial effect after taking sickness and complications not immediately fatal into account. (We did not count these complications if the patient died that day, but we did if they survived that day and died later.) Thus, death may be cutting off costs of patients who would be expected from initial sickness or complications to be very expensive.

Complications not immediately fatal have an enormous impact on costs: The complications scales are 0 for those with no in-hospital complications and have been rescaled here so that a 1 percentage point rise in the complications scale is associated with a 1 percentage point increase in death at 180 days (after adjusting for sickness at admission).

In results not shown here, we regressed log (standardized costs) on sickness and process but not in-hospital outcomes. This allowed us to judge how much the cost

Table 3.7

REGRESSION OF LOG (STANDARDIZED COSTS) ON SICKNESS, PROCESS, AND OUTCOMES

	CHF		AMI		PNE		CVA		HIP	
Variables	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
In-hospital death	−0.234	−3.76	−0.352	−9.72	−0.603	−10.73	−0.429	−8.56	−0.261	−3.48
Complications	3.24	20.71	2.86	22.50	3.23	19.26	3.88	19.54	2.22	15.94
Sickness	0.532	7.54	0.454	6.57	0.506	8.40	0.596	9.31	0.646	8.86
Process	0.091	4.95	0.114	7.10	0.092	4.92	0.139	7.27	0.048	3.83
Constant	−0.268	−13.46	−0.230	−12.13	−0.195	−9.19	−0.181	−8.72	−0.106	−8.11
N	1276		1222		1227		1246		1227	
R-squared	0.36		0.42		0.37		0.37		0.29	

impacts of these variables are driven by their links to unfortunate in-hospital outcomes. The process coefficients did not change; indeed, we know from work published elsewhere that process is associated with better, not worse, in-hospital outcomes. By contrast, the sickness coefficients were much smaller when we included bad in-hospital outcomes, implying that the impact of sickness in increasing expensive nonfatal complications overshadows any effects on increasing death.

EFFECTS OF LENGTH OF STAY ON COSTS

To determine how other factors affect the cost per day, we could adjust log (standardized costs) for log (length of stay). People with one-day lengths of stay who were discharged alive had costs much higher than expected—they are probably people who transferred out. In our data processing of the Bill Retrieval File, we would add to get the total bill for all parts of the hospitalization episode and not parcel it out. For this analysis, we deleted all such people (three for CHF, 15 for AMI, four for CVA, zero for PNE, one for HIP). Because the cost per day was a little higher for shorter stays, we used log (length of stay + 1) as the explanatory variable. The added "one day" might reflect fixed costs of a hospitalization (initial tests, etc., that would be done no matter how long the stay).

Naturally, as Table 3.8 shows, length of stay is enormously powerful as an explanatory variable. Since the coefficient of log (length of stay + 1) is close to 1 as we expected, the coefficients of the other variables reflect their impact on the cost per day rather than total costs as they did earlier. Comparing the coefficients in Table 3.8 with those in Table 3.7, we see that in-hospital death increases costs per day and the effects of complications and sickness are greatly reduced (i.e., much of their earlier effect on total costs was due to their effect on length of stay). The effects of process are about the same, showing that better process is associated with higher costs per day but not longer stays.

Table 3.8

REGRESSION OF LOG (STANDARDIZED COSTS) ON LENGTH OF STAY, SICKNESS, PROCESS, AND OUTCOMES

	CHF		AMI		PNE		CVA		HIP	
Variables	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Log (los + 1)	0.929	39.55	0.787	34.97	1.005	40.85	0.995	48.63	0.745	35.18
In-hospital death	0.142	3.31	0.385	11.61	0.171	4.16	0.284	8.66	0.249	4.54
Complications	1.04	8.76	1.10	10.70	0.76	6.14	1.06	8.13	0.96	9.25
Sickness	0.369	7.78	0.126	2.55	0.263	6.64	0.255	6.67	0.289	5.52
Process	0.078	6.32	0.095	8.43	0.092	7.62	0.104	9.26	0.032	3.65
Constant	–2.118	–43.55	–1.995	–38.21	–2.36	–43.09	–2.431	–50.81	–2.021	–36.62
N	1276		1222		1227		1246		1227	
R-squared	0.71		0.71		0.73		0.78		0.64	

IV. THE DISTRIBUTION AND DETERMINANTS OF COSTS ACROSS HOSPITALS

RELATION OF COSTS TO PAYMENT VARIABLES

We first examined how the relationship between hospital costs and the payment factors used by HCFA is affected by patient sickness. To do so, we regressed the log of costs on the payment variables and the cost-sickness scales. At the top of Table 4.1, the coefficients are shown when no adjustment is made for sickness; at the bottom, the regressions with the cost-sickness scales included are shown.

The coefficients on the payment variables are close to what they should be in principle[1] when sickness is not included but are a little smaller when sickness is included. They are smallest for pneumonia, where sickness has the largest effect on costs.

Table 4.1

REGRESSIONS OF LOG (COSTS) ON PAYMENT AND SICKNESS VARIABLES

	CHF	AMI	PNE	CVA	HIP
Payment Variables					
Log (DRG weight)	1.00	1.12	0.77	1.17	1.31
Disp. share	0.85	0.31	0.53	0.92	1.39
Log (wage index)	1.19	1.49	1.25	1.24	1.10
Irb	0.19	0.67	0.10	0.71	0.35
Constant	12.42	12.42	11.98	12.23	12.15
R-squared	0.16	0.34	0.09	0.12	0.30
N	1279	1237	1227	1250	1228
Payment Variables with Sickness Included					
Log (DRG weight)	0.94	1.06	0.52	0.95	1.17
Disp. share	0.88	0.19	0.80	0.34	1.26
Log (wage index)	1.11	1.42	0.93	1.25	1.10
Irb	0.12	0.61	-0.26	0.71	0.25
Sickness scale	1.01	1.00	1.02	1.01	1.01
R-squared	0.25	0.41	0.25	0.25	0.38

[1]In principle, if the DRG weights are correct, the coefficients on log (DRG) should be 1. The coefficient on log (fiscal wage index) should be 1/0.79. (For x small, log [a + (1 − a)(1 + x)] ≃ (1 − a) log (1 + x).)

In particular, including the sickness scale reduces the association of the DRG weights with costs. Thus, the same sickness variables associated with higher costs within DRGs are also associated with higher-weight DRGs. Adding in the sickness variables reduces the compression of the DRG weights noted elsewhere as being associated with bigger hospitals (Thorpe, Cretin, and Keeler, 1988). Controlling for sickness also reduces the effects of hospital teaching status on costs but has mixed effects on disproportionate share and on the wage index.

RELATION OF STANDARDIZED COSTS TO OTHER HOSPITAL FACTORS

In the rest of this section, we will analyze average hospital values of standardized costs and other factors. In particular, we averaged each variable over the approximately 4.6 patients per disease, and after setting means to zero for each disease by subtracting the mean, averaged the five diseases for each hospital. Means and standard deviations of these hospital-level variables are given in Table 4.2. The distributions of other hospital variables used in the analysis (number of beds, type of control, SMSA size or rural location, state, percentage Medicaid, and percentage Medicare) are given in Draper et al. (forthcoming) (or see Table 5.3).

Hospital differences are much greater in profitability than in sickness. The standard deviation of the hospital component of sickness is 0.069; of log (standardized costs), 0.26. Both are scaled by their effects on log (standardized costs), so the sickness differences can be described as 0.069/0.26 = 27 percent as big as cost differences.

Table 4.2

HOSPITAL AVERAGE PATIENT-LEVEL VARIABLES

Variable	Mean	Standard Deviation
Log (standardized costs)	0	0.26
Standardized costs[a]	1.256	0.39
Cost-sickness scale values	0	0.069
In-hospital mortality	0.128	0.082
In-hospital mortality (within 5 days postadmission)	0.066	0.064
In-hospital complications	0.073	0.029

[a]Calculated by exponentiating (the centered) log (standardized costs).

A component of variance analysis described in Appendix A attempts to remove the upward bias in these numbers from our sample of only five diseases. It indicates that hospitals have a standard deviation of 0.21 in log (standardized costs), but only 0.045 in sickness.

WHICH HOSPITALS HAD HIGHER PROFITS, CONTROLLING FOR SICKNESS?

We looked at other hospital-level variables that might be associated with average standardized costs and average sickness (see Table 4.3). Teaching and percentage Medicaid had little effect, undoubtedly because they are already taken into account in the payment formula. We were not able to explain a high percentage of the variance with the hospital-level variables we tried to use (R-squared = 0.25). Average sickness is the most powerful predictor of loss.[2] State effects were fairly small, although one state had 6 percent more profit than the others. City size as shown by a six-way classification of

Table 4.3

REGRESSION OF AVERAGE LOG (STANDARDIZED COSTS) ON HOSPITAL VARIABLES

	Coef.	t	Coef.	t
Sickness	1.37	6.7	NA	NA
State4	−0.06	−1.8	−0.04	−1.2
>400 beds	0.04	1.3	0.06	1.9
Urban government control	−0.19	−5.1	−0.20	−5.0
>40% Medicare admissions	0.06	1.7	0.04	1.2
Rural	−0.06	−1.8	−0.13	−3.5
R-squared	0.25		0.14	
Number of hospitals	296		296	

[2]The sickness scales were fit on the data sets used to estimate the variance they explain. To the extent that data were overfit, the variance estimates presented in this section will be too high. To judge the amount of overfitting in defining the sickness scale that predicts costs, we computed the R-squared of an ordinary least-squares regression of log (standardized costs) on the sickness variables and contrasted it with a conservative estimate (see Sec. II) of R-squared based on the same variables. The results in the table show this to be a minor problem.

	CHF	AMI	PNE	CVA	HIP
Regular R-squared	0.115	0.119	0.143	0.148	0.127
Conservative R-squared	0.103	0.100	0.109	0.126	0.100

SMSA size had little effect, although rural hospitals were slightly more profitable. Larger hospitals were slightly less profitable. Urban-government hospitals had much less loss per case than either nonprofit or for-profit hospitals. Hospitals with a high percentage of Medicare patients had higher loss per Medicare case. If sickness is not included (last two columns of Table 4.3), large hospitals do worse and rural hospitals do better, indicating the presence of more expensive patients at large hospitals and less expensive patients at rural hospitals.

WHAT PROPORTION OF HOSPITAL AVERAGE COSTS IS JUSTIFIABLE?

In the Medicare PPS, payments to hospitals for an admission are based on the national average cost per case for that diagnosis. Even after adjustment for the payment factors, enormous variations in average hospital costs remain, and since 1988, hospital-specific costs have not affected payments.

Recently, many health economists (Goodall, forthcoming; Pope, forthcoming; Lave, forthcoming) have argued that the transition from hospital-specific costs to adjusted national average costs has gone too far.[3] They say it would be better to pay a blend of hospital-specific and adjusted national costs (as was done in the PPS phase-in period from 1984–1988). They argue that (1) the massive transfers from high- to low-cost hospitals are not necessary for cost-control incentives to operate and may well have harmful financial impacts, and that (2) some portion of the cost differences is justified.

Such equity arguments for blended cost payments complement efficiency arguments made previously. Ellis and McGuire (1986) showed that if providers value their own profits more than benefits to patients, then hospitals paying all the marginal costs of care may underserve patients.[4] Ellis and McGuire also note that blending will mitigate other possible problems of prospective payment: access difficulties for predictably expensive patients, inappropriate admission of cheap patients, and excessive hospital competition for cheap doctors. Newhouse (1990) argues that blending will help in any situation in which administered prices may be incorrect.

---

[3]O'Day and Dobson (forthcoming) strongly oppose this argument.

[4]The RAND PPS/quality-of-care study found little evidence of such underservice: Indeed, we found few noticeable effects of PPS on trends over time to higher quality of care in hospitals (Rogers et al., forthcoming). However, the "after PPS" data in that study come from admissions between July 1985 and June 1986, when a transitional 50–50 blend of national and hospital-specific costs was in effect.

Goodall (forthcoming) and Pope (forthcoming) both show under certain assumptions that average squared unfairness is minimized when the weight on hospital-specific costs in the blend is the regression coefficient of hospital average costs on justified costs. If the hospital average justified and unjustified costs are uncorrelated, this is equal to the R-squared of a regression of average costs on justifiable variables. Such a blended formula makes sense only if data on unjustifiable factors are too expensive to collect; using the hospital blend is optimal under this presumption.[5] Justifiable variables that, like the current payment adjustments, could easily be collected at the hospital level should be used to adjust payments directly and would not affect the blending proportion.[6]

WHICH CURRENTLY UNCOMPENSATED COST FACTORS ARE JUSTIFIABLE?

The least controversial factor would be patient sickness at admission, which we have measured here directly. Other measures highly related to sickness and costs include quick death and in-hospital complications.

At the individual-patient level, quick death leads to lower costs, but later in-hospital death and in-hospital complications lead to higher costs. Here, we will only compute their effects on average costs, and, hence, the share of hospital-specific costs in a blended-payment formula. However, we will comment briefly on an alternative use: as a direct adjustor of payments to a hospital. Quick death might be suitable for that purpose, since the financial incentives would lead to keeping the patient alive and recording would not be a problem[7] (one might want a sliding scale to prevent hospitals from dragging an essentially dead patient over the "quick" limit). Most quick deaths are not preventable (the physicians in Dubois et al. (1987) estimated that 75–85 percent of the in-hospital deaths in that sample were not preventable, depending on the stringency of the definition). However, nonfatal complications are associated with higher payments leading to the wrong incentives, and in addition might be more susceptible to fudging in the charts.

---

[5]If collecting data on a justifiable factor were easy and cheap, using those data to adjust prospective payments directly would be better. This would be fair and would have no adverse effects on incentives. Unfortunately, most patient-level data are not cheap to collect.

[6]When average standardized costs are regressed on payment variables and sickness, no payment variable is close to significant ($t < 0.5$). Hence, we will analyze standardized costs, assuming that the proposed payment would blend the current prospective payments with hospital-specific costs.

[7]Indeed, death is currently used in the DRG system, with AMI discharged dead receiving a lower payment than AMI discharged alive.

Although we have not done so here, one might want to include other factors that add to expense as justifiable: cost-raising problems such as small size or isolation, difficulty of treating Medicare and other patients (i.e., different languages, psychiatric disorders, alcoholism, and difficulties in placing convalescent patients).

RESULTS

We regressed the average (raw, not logged) standardized (for payment factors) costs per case on averaged (over the same 23 patients per hospital whose costs were averaged) patient variables. Table 4.4 gives the resulting R-squared for various combinations of variables measuring sickness, complications, and death. In all cases, the regression is based on 296 hospitals. As expected, average levels of sickness and nonfatal complications substantially increased average costs, and quick death reduced costs. When quick death and in-hospital death were both used as predictors, later in-hospital death added substantially to costs.

In these data, depending on what is considered justifiable, 15–29 percent of the variance in average standardized hospital costs is explained by justifiable factors.[8] Hence, according to the equity arguments of Pope (forthcoming) and Goodall (forthcoming), 15–29 percent is a lower bound on the desirable weight on hospital-specific costs in an blend. Some of the remaining variance is noise, some comes from

Table 4.4

REGRESSION OF HOSPITAL AVERAGE STANDARDIZED COSTS ON SEVERITY VARIABLES
(In thousands of U.S. dollars)

	Coef.	t	Coef.	t	Coef.	t	Coef.	t
Sickness	5.2	7.2	5.3	6.7	3.3	4.4	4.5	6.2
Death in 5 days	NA	NA	–2.6	–3.0	–2.6	–3.5	–6.3	–6.2
Nonfatal complications	NA	NA	NA	NA	1.2	6.6	NA	NA
In-hospital death	NA	NA	NA	NA	NA	NA	4.4	5.5
R-squared	0.15		0.18		0.29		0.26	

[8]If we did not want to pay hospitals more for inefficiency or more intense treatment, we should control for these factors, but unfortunately we did not have good proxies for them. Here, we are simply assuming that justifiable and unjustifiable costs are uncorrelated.

differences in treatment intensity and efficiency, and some from justifiable factors that we did not measure in the study.[9]

A blend may be more equitable but would put less pressure on cost control at hospitals than the current system (and even if historic costs are used for the hospital-specific portion of payments, hospitals will get a portion of current expenditures back in the future if the costs are updated). Despite the costs of hospital care, relaxing this cost-control pressure on hospitals is desirable from an efficiency standpoint as well. Especially as more kinds of payments for health care become prospective, giving providers some financial incentives for supplying expensive care when it is needed will be helpful.

[9]These estimates are limited to our sample of five diseases, which, although common, tended to be more serious and emergent than average. A more general sample of diseases might lead to different results.

V. DETERMINANTS OF OUTLIERS

In this section, we will try to see whether the extremely expensive outlier cases are caused by the same factors as expensive cases generally or whether they represent a radically different phenomenon. Because of the coding problems we discuss below, the analyses are based on synthetic outliers–cases that because of their extreme costs should have received extra payments according to the regulations. Only two-thirds of these cases were coded as actually having received outlier payments.

CASES THAT SHOULD HAVE RECEIVED PAYMENTS BUT DID NOT

Day and cost are the two types of outliers. In each DRG, both day outliers and cost outliers have thresholds. In principle, the hospital would get additional outlier payments for cases that exceeded either of these thresholds. The cost threshold was the maximum (12,500, 2 * (DRG payment)). The DRG payment is the product of the weight for the DRG and the hospital payment per unit weight, which is computed for each hospital based on several factors. For example, for DRG 121, "circulatory disorders with AMI and complications, discharged alive," the FY 1986 threshold is 27 days and the DRG weight is approximately 1.77.

Many cases that should have been cost outliers according to the regulations are not coded as being so. Either the intermediary did not make the outlier payments or made them and did not code them properly. In the Bill Retrieval File we used for this analysis, two items refer to outliers, outlier payments made, and a code variable that indicates "not an outlier," "day outlier," or "cost outlier." On cases showing outlier payments, but with the code variable indicating "not an outlier," we assume that the code variable is incorrect. (In only one case did the code variable indicate outlier with payments shown as zero.) There were half as many cases with very large charges and getting no payments as there were cases shown with payments (see Table 5.1).

In what follows we will define synthetic outliers to include all cases that received payments, plus all those that should have (to be conservative, we tightened the requirement to costs greater than three times the DRG payments rather than two). Synthetic outliers constituted 4.5 percent of the cases in our sample.

Table 5.1

NUMBER OF CASES WITH OUTLIER PAYMENTS

Outlier Type	CHF	AMI	PNE	CVA	HIP	Total
Day	8	8	19[a]	26	15	76
Cost	12	28	17	23	28	108
Not coded	5	7	6	5	1	24
Total paid	25	43	41	54	44	207
Should be cost[b]	19	17	23	21	19	99
Total cases	1279	1237	1227	1250	1228	6251

[a]Includes one day outlier with payments = 0.
[b]Included if costs are greater than two times the DRG payments and if charges are greater than $20,000.

HOW ARE OUTLIERS DIFFERENT?

In Sec. III, we determined factors that lead to higher costs generally. We will now check whether these factors extend smoothly to the most expensive cases. Table 5.2 compares the means of many variables for (synthetic) outliers with those of a 30 percent random sample of nonoutliers. Table 5.3 gives cross-tabulations of synthetic outliers by interesting discrete variables. Outlier cases do not differ much by age, sex, Medicaid status, or instability at discharge. Outlier cases are somewhat sicker at admission, but differ most in in-hospital complications and in DNR after day 1. A DNR at day 1 is much more common for nonoutliers. Process, especially the ICU, physician cognitive scale, and technical diagnostic scale, is higher on average for outliers. The average DRG weight for outliers is only slightly higher than for nonoutliers, but outliers are more likely to come from bigger, teaching, and nonrural hospitals. Table 5.4 shows the results of weighted logistics regressions of (synthetic) outlier status on important independent variables. The main qualitative difference from earlier results on costs generally is that in-hospital death is not (negatively) associated with outlier status, although it is associated with lower costs generally. The statistical significance of the results is much less because outlier cases are so rare.

Table 5.2

MEANS OF PROCESS, COST, SICKNESS, AND OUTCOME VARIABLES BY OUTLIER STATUS

	Synthetic Nonoutliers	Synthetic Outliers[a]	Nonoutliers' Standard Deviation
Hospital Characteristics			
Irb	0.060	0.116	0.130
Beds	294.052	371.022	218.783
City-county	0.115	0.202	0.319
High Medicaid hospitals	0.180	0.232	0.384
Rural	0.223	0.070	0.416
Cost/charge ratio	0.646	0.580	0.130
Disproportionate share	0.027	0.043	0.047
Process Variables			
ICU	–0.038	0.431	0.668
Physician cognitive	0.135	0.435	0.941
Nurse cognitive	0.180	0.125	0.902
Technical/therapeutic	0.086	0.070	0.940
Technical/diagnostic	0.094	0.386	0.959
Total process	0.199	0.560	0.926
Cost Variables			
Charges ($)	6596	33,312	4953
Outlier payments	0	1,153	0
Number of surgeries	0.954	1.938	1.117
DRG weight 86	1.505	1.640	0.599
Log (cost)	12.692	14.341	0.700
Sickness[b]	–0.011	0.215	0.235
Log (standardized costs)	–0.050	1.450	0.594
Standardized costs	1.128	4.768	0.699
Patient Variables			
Age	78.186	77.665	7.663
Male	0.444	0.474	0.497
Length of stay	10.196	29.180	5.708
Medicaid	0.135	0.173	0.341
Died in hospital	0.109	0.301	0.312
Die (30 days)	0.147	0.239	0.355
Die (180 days)	0.280	0.522	0.449
DNR day 1	0.252	0.136	0.435
Later DNR	0.055	0.294	0.228
Discharge sickness[c]	0.928	1.163	1.080
Discharge instability[c]	0.342	0.421	0.686
Complications[d]	0.067	0.272	0.102
Sickness at admission (30)	0.153	0.259	0.174
Sickness at admission (180)[d]	0.288	0.419	0.223
Number of observations	1703[e]	272	NA

[a]Cases were called synthetic outliers if costs/payments greater than three or outlier payments were reported. Nonoutliers numbers are based on a 30 percent sample of nonoutliers for each disease.

[b]Sickness scale that best predicts costs.

[c]If discharged alive.

[d]Complications and sickness are expressed in terms of 180-day postadmission mortality.

[e]A 30 percent sample of nonoutliers in each disease.

Table 5.3

CHARACTERISTICS OF PATIENTS AND HOSPITALS BY OUTLIER STATUS

	Synthetic Nonoutliers	Synthetic Outliers[a]	Total Patients
Ownership			
Nonprofit	1016	181	1197
Profit	227	48	275
Government	460	43	503
Teaching Status			
Nonteaching	1151	144	1295
Minor	407	71	478
Major	145	57	202
States			
A	371	40	411
B	339	73	412
C	310	66	376
D	363	66	429
E	320	27	347
Outlier Code			
Day outlier	0	76	76
Cost outlier	0	108	108
Payments[b]	0	25[c]	25
No payments	1703	63	1766
Patient Origin			
Home	1382	225	1607
Nursing home	206	27	233
Other	59	9	68
No data	56	11	67
Disease			
HIP	333	52	385
AMI	354	52	406
CHF	350	41	391
PNE	347	58	405
CVA	319	69	388
Destination of Patient			
Home	971	84	1055
Nursing home	384	75	459
Residential home	28	8	36
Hospital	115	22	137
No data	19	1	20
Dead at discharge	186	82	268
Total	1703	272	1975

[a]Cases were called synthetic outliers if costs/payments greater than three or outlier payments were reported. Nonoutliers numbers are based on a 30 percent sample of nonoutliers for each disease.

[b]Coded as not outlier, but some outlier payments made.

[c]Includes one with code = 7, presumably an error in the Bill Retrieval File.

Table 5.4

REGRESSION OF OUTLIER STATUS ON SICKNESS, PROCESS, AND OUTCOMES

	Coef.	t	Coef.	t
Died in hospital	–0.37	–1.09	1.40	3.23
Complications	6.98	8.94	3.55	3.66
Sickness[a]	2.07	4.47	1.70	2.88
Process	0.17	1.16	0.35	2.01
Log (length of stay + 1)	NA	NA	4.48	11.13
Constant	–4.20	NA	–16.85	NA
Number of observations	1975[b]		1975[b]	
Chi-square	174		396	

[a]If discharged alive.
[b]Regression is weighted to reflect 30 percent sampling of nonoutliers.

Appendix A

VARIANCE COMPONENT CALCULATIONS OF HOSPITAL COSTS AND SICKNESS

We wanted to know how sickness and loss per case are distributed across hospitals. We used the same components-of-variance model to analyze each. In the body of the Note, we simply averaged all observations at each hospital. This method includes some disease-hospital effects as well as hospital effects and so overestimates the hospital-specific components that would be obtained if we included all diseases (not just our five).

Let $X(h, d, i)$ be the value of either sickness, the cost-sickness scale, or log (loss)—the log (adjusted costs)—for patient i with disease d at hospital h. Our sample has a little more than four patients on average in each hospital disease cell. Assume

$$X(d, h, i) = a(h) + b(d, h) + c(d) + e(d, h, i),$$

where $a(h)$ is the underlying true hospital average with mean 0, $b(d, h)$ is the underlying true hospital-disease average with mean 0, $c(d)$ is the underlying true average value for the disease, and $e(d, h, i)$ is independent random noise.

We have subtracted the average X for each disease to make $c(d) = 0$ in all cases. For losses, this avoids the issue of which disease is more profitable. Also, sickness has no inherent zero and has also been centered. Sickness is scaled so that a 1 percent change leads to a 1 percent change in profitability.

We will estimate the variance of each component of the model. If hospitals vary consistently, the variance of a and b will be large.

We first performed an analysis of variance on sickness and log (loss) for all diseases combined (see Table A.1). Next, we performed a similar analysis on the losses and sickness for each disease separately to estimate the hospital-disease component b (see Table A.2). Using the 5 disease mean as the estimate for b, we make the inferences shown in Table A.3.

Hospital differences are much greater in profitability than in sickness. Taking the square root of the variance of a, we obtain the standard deviation of the hospital

Table A.1

ANALYSIS OF VARIANCE (ALL DISEASES COMBINED)

	Components of Variance		
	Total	*a*	*b + e*
Sickness	0.249	0.0020	0.247
Log (loss)	0.450	0.0430	0.407

Table A.2

ANALYSIS OF VARIANCE (EACH DISEASE)

	Components of Variance	
	Sickness	Log (loss)
Disease	*(a + b)*	*(a + b)*
HIP	0.0012	0.041
AMI	0.0017	0.058
CVA	0.0018	0.061
CHF	0.0027	0.051
PNE	0.0107	0.066
Mean of 5 diseases	0.0036	0.056

Table A.3

ANALYSIS OF VARIANCE (FINAL ESTIMATES)

	Components of Variance			
	Total	*a*	*b*	*e*
Sickness	0.249	0.0020	0.0016	0.245
Log (loss)	0.450	0.0430	0.013	0.394

component of sickness (0.045) and of losses (0.207). The sickness measure is scaled by its effects on loss so the measures can be directly compared. Only 20 percent of the loss differences are due to differences in sickness at admission.

Appendix B

EFFECTS OF OUTLIER PAYMENTS ON HOSPITAL BEHAVIOR

Hospitals receive outlier payments for extremely expensive stays, in effect insuring them against a portion of their risk (Ellis and McGuire, 1988). They have the form of a very large deductible: In return for the fixed DRG-based payment, the hospital is responsible for the first N days of care (in the case of day outliers) or the first M dollars (in the case of cost outliers), and then HCFA pays much of the rest of the costs of care. Assuming that hospitals do not change their behavior as a result of outlier payments, such a payment structure is the best that can be done case by case to minimize financial risk to the hospitals when the total amount of additional payments is limited (Keeler, Carter, and Trude, 1988).

However, the outlier payments themselves may induce inappropriate levels of care—too much care for the outlying patient, for whom the hospital is reimbursed at the margin, and too little care for the expensive patient who is not an outlier, for whom the hospital gets nothing at the margin. In addition, at least some expensive cases are due to adverse events resulting from hospital errors (Haley et al., 1987; Zook and Moore, 1980). If the expensive cases result from hospital inefficiency, then perhaps hospitals should not get extra payments for them. If, on the other hand, the expensive cases result from sicker patients, then hospitals who take more than their share of such patients may be underpaid. Recent attempts to adjust DRGs to better reflect patient sickness might help this problem.

In principle, we could look at outliers to study both the determinants of hospital expense (in extreme form) and the results of HCFA's cost sharing on hospital behavior.

We originally planned to study whether hospitals took outlier payments into account in their decisions on treatment and discharge of such patients. Unfortunately, the data did not permit us to address this question in a satisfying way. There were two main problems. First, from July 1985 to June 1986, many cases that should have been cost outliers were not coded as such in the Bill Retrieval File. From 1985 to early 1986, although day outliers were always computed, the hospitals had to request reimbursement for cost outliers. If the hospitals were not savvy enough to request the outliers, the

outlier policy would not likely have any effect on their behavior. Second, because the fixed-cost system was only 50 percent of payments in this period, the incentives posed by outliers were weak. As a result, although our preliminary investigations showed little effect of outlier payments on behavior, we do not know whether this means that outlier incentives do not influence behavior today or whether the noise from coding and payment errors, and the weakness of the early incentives, prevented us from discovering a true effect.

WEAKNESS OF INCENTIVES

To compute the day outlier payments for this period, we (1) find the daily rate—(hospital payment for that DRG)/geometric mean length of stay—and (2) outlier payments—(length-of-stay threshold) $\times$ 0.6 $\times$ 0.5 $\times$ daily rate. The 0.6 is the coinsurance factor; 0.5 was used because the prospective portion of payments in 1985 and early 1986 was still only 50 percent. The Prospective Payment System was supposed to be more completely phased in (i.e., the proportion of payments based on DRGs as opposed to historic costs was supposed to rise) in FY 1986, but in fact the increase did not occur until May 1, 1986.

To compute cost outlier payments, we took the total charges $\times$ 0.66 (the national average cost-to-charge ratio), subtracted the maximum ($12,500, or 2 $\times$ the DRG payment), and again multiplied by 0.6 $\times$ 0.5. In these data, most outliers were cost outliers (perhaps because of the seriousness of the diseases). For these conditions, the cost outlier payments are only 30 percent of the marginal costs of care. For day outliers, payments might be a little higher than 30 percent of marginal costs if the average cost of outlier days was less than the normal daily rate. Thus, even for cases on which outlier payments were received, the payments were very small.

Table B.1 shows that the average outlier payments for each disease ranged from $1200–$2000 and were about 5 percent of the total charges for the cases that were paid. Note also that the outlier cases in the table are only slightly more expensive than the "should be" outliers. Indeed, the main reason for any difference is the $1 or $2 million cases in each disease, which were always marked as outliers.

Table B.1

CHARACTERISTICS OF OUTLIER AND "SHOULD-BE-OUTLIER" CASES

	Number	LOS (Days)	Total Charge	Outlier Payments	Log (Adj. Costs)[a]
CHF outlier	25	29	26,458	1332	1.65
Should be[b]	19	21	24,951	0	1.52
AMI outlier	43	26	38,428	1768	1.23
Should be[b]	17	26	38,737	0	1.30
PNE outlier	41	30	35,413	2050	1.53
Should be[b]	23	25	29,372	0	1.31
CVA outlier	54	31	31,492	1250	1.56
Should be[b]	21	26	28,191	0	1.27
HIP outlier	44	33	32,356	1203	1.16
Should be[b]	19	33	25,874	0	1.12

[a]After subtracting the mean for all cases.
[b]Included if costs are greater than two times the DRG payments and if charges are greater than $20,000.

EFFECTS OF OUTLIER PAYMENTS ON BEHAVIOR

This analysis was supposed to compare the instability at discharge of outlier cases and "near-outlier" cases (that ended up a little cheaper or shorter than the thresholds). The hypothesis was that outlier payments might allow hospitals to keep expensive patients until they got better or died. In this case, we would expect near-outlier patients to be more unstable at discharge and less likely to die in the hospital than outlier cases. We would compare in-hospital death, death within a few days of discharge, and instability at discharge of live discharges. In fact, we will define three groups: cases with outlier payments, cases we think should have received outlier payments, and cases that were near but below the thresholds. The nonrecognized outlier cases should be an even better control group than the near outliers for any behavioral effects.

For AMI, the only disease tried, instability at discharge for live discharges did not vary across the three groups, but in-hospital death was somewhat higher for cost outliers than for the "near-cost outliers" and "should-have-been" outliers, as expected.

BIBLIOGRAPHY

Daley, J., S. F. Jencks, D. C. Draper, et al., "Predicting Hospital-Associated Mortality for Medicare Patients," *JAMA*, 1988;260;24:3617–3624.

Draper, D. C., K. L. Kahn, E. J. Reinisch, et al., "Studying the Effects of the DRG-Based Prospective Payment System on Quality of Care: Design, Sampling, and Fieldwork," *JAMA*, forthcoming.

Dubois, R. W., W. H. Rogers, J. H. Moxley, et al., "Hospital Inpatient Mortality: Is It a Predictor of Quality," *New England Journal of Medicine*, 1987:317;1674–1680.

Ellis, R. D., and T. G. McGuire, "Insurance Principles and the Design of Prospective Payments Systems," *J. Health Econ.*, 1988.

Ellis, R. D., and T. G. McGuire, "Provider Behavior under Prospective Reimbursement: Cost Sharing and Supply," *J. Health Econ.*, 1986;5:129–151.

Goldman, L., D. L. Caldera, S. R. Nussbaum, et al., "Multifactorial Index of Cardiac Risk in Noncardiac Surgical Procedures," *New England Journal of Medicine*, 1977;297;845–850.

Goodall, C. R., "A Simple Method for Determining Standard Percentage in Reimbursement Applications," *JHE*, forthcoming.

Guterman, S., P. W. Eggars, G. Riley, et al., "First Three Years of Medical Prospective Payment: An Overview," *HCFA Review*, 1988:67–77.

Haley, R. W., J. W. White, D. H. Culver, and J. M. Hughes, "The Financial Incentives for Hospitals to Prevent Nosocomial Infections under the Prospective Payment System," *JAMA*, 1987;257:1611–1614

Health Care Financing Administration, *Medicare Hospital Mortality Information*, HCFA Publication No. 00642, Washington, D.C., 1987.

Jencks, S. F., and T. L. Kay, "Do Frail, Disabled, Poor, and Very Old Medicare Beneficiaries Have Higher Hospital Charges?" *JAMA*, 1987;257:198–202.

Jencks, S. F., D. K. Williams, and T. L. Kay, "Assessing Hospital Associated Deaths from Discharge Data—The Role of Length of Stay and Comorbidities," *JAMA*, 1988;260:2240–2246.

Kahn, K. L., R. H. Brook, D. C. Draper, et al., "Interpreting Hospital Mortality Data," *JAMA*, 1988a;260;3625–3628.

Kahn, K. L., M. R. Chassin, L. V. Rubenstein, et al., *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients with Congestive Heart Failure*, The RAND Corporation, N-2798-HCFA, December 1988b.

Kahn, K. L., W. H. Rogers, and L. V. Rubenstein, "Measuring Quality of Care with Explicit Process Criteria Pre- and Post-Implementation of the DRG-Based Prospective Payment System," *JAMA*, forthcoming.

Keeler, E. B., G. M. Carter, and S. Trude, "Insurance Aspects of DRG Outlier Payments," *J. Health Econ.*, 1988.

Keeler, E. B., K. L. Kahn, D. C. Draper, et al., "Changes in Sickness at Admission Following the Introduction of the Prospective Payment System," *JAMA*, forthcoming.

Killip, T., and J. T. Kimball, "Treatment of Myocardial Infarction in a Coronary Care Unit," *Am. J. Cardiol.*, 1967;20:457–461.

Knaus, W. A., E. A. Draper, D. P. Wagner, et al., "APACHE II: A Severity of Disease Classification System," *Crit. Care Med.*, 1985;13:818–829.

Kosecoff, J., K. L. Kahn, W. H. Rogers, et al., "Prospective Payment System and Impairment at Discharge: The Quicker-and-Sicker Story Revisited," *JAMA*, forthcoming.

Kosecoff, J., L. V. Rubenstein, K. L. Kahn, et al., *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients with Acute Myocardial Infarction*, The RAND Corporation, N-2799-HCFA, December 1988.

Lave, J. R., *Yale Journal of Regulation*, forthcoming.

Newhouse, J. P., "Pricing Behavior and Imperfections in the Medical Care Marketplace," talk given at the 1990 World Conference of Health Economics, September 1990.

Norris, R. M., P.W.T. Brandt, D. E. Caughey, et al., "A New Coronary Prognostic Index," *The Lancet*, 1969;1:274–278.

O'Day, L., and A. Dobson, *Yale Journal of Regulation*, forthcoming.

Pope, G. C., "Using Hospital-Specific Costs to Improve the Fairness of Prospective Reimbursement," *J. Health Econ.*, forthcoming.

Rogers, W. H., D. C. Draper, K. L. Kahn, et al., "Quality of Care Before and After Implementation of the DRG-Based Prospective Payment System: A Summary of Effects," forthcoming.

Roth, C. P., K. L. Kahn, M. J. Sherwood, et al., *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients with Pneumonia*, The RAND Corporation, N-2801-HCFA, December 1988.

Rubenstein, L. V., J. Kosecoff, K. L. Kahn, et al., *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients with Cerebrovascular Accident*, The RAND Corporation, N-2802-HCFA, December 1988.

Schroeder, S. A., J. A. Showstack, and H. E. Roberts, "Frequency and Clinical Description of High-Cost Patients in 17 Acute-Care Hospitals," *New England Journal of Medicine*, 1979;300:1306–1310.

Schroeder, S. A., J. A. Showstack, and J. Schwartz, "Survival of Adult High-Cost Patients: Report of a Follow-Up Study from Nine Acute-Care Hospitals," *JAMA*, 1981;245:1446–1449.

Sherwood, M. J., K. L. Kahn, J. Kosecoff, et al., *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients with Hip Fracture*, The RAND Corporation, N-2800-HCFA, December 1988.

Thorpe, K. E., S. Cretin, and E. B. Keeler, "Are the Diagnosis-Related Group Case Weights Compressed," *HCFA Review*, Winter;1988:37–46.

Trude, S., and G. M. Carter, *A Description of Expensive and Long-Staying Patients*, The RAND Corporation, N-2998-HCFA, October 1989.

Zook, C. J., and F. D. Moore, "High Cost Users of Medical Care," *New England Journal of Medicine*, 1980;302:996–1002.