EDITORIAL

WHAT PROPORTION OF HOSPITAL COST DIFFERENCES IS JUSTIFIABLE?

Emmett B. KEELER
RAND, Santa Monica, CA 90406-2138, USA

Final version received June 1990

1. Introduction

In the Medicare Prospective Payment System (PPS), payments to hospitals for an admission are based on the national average cost per case for that diagnosis. The payments are adjusted for local labor costs, urban–rural location, treatment of a disproportionate share of Medicaid patients, amount of teaching, and outlier payments for unusually expensive or lengthy cases. Still, enormous variations in average hospital costs remain, and since 1988 hospital-specific costs have not been part of the payment formula.

Two papers [Goodall (1990), Pope (1990)] in this issue argue that the transition from hospital-specific costs to adjusted national average costs has gone too far. 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:

(1) that the massive transfers from high-cost to low-cost hospitals are not necessary for cost-control incentives to operate, and may well have harmful financial impacts; and
(2) some portion of the cost differences are justified.

These equity arguments for blended cost payments complement efficiency arguments made previously. With fully prospective payment, hospital payments for a Medicare patient are the same no matter what care they provide, so hospitals are paying all the marginal costs of care. Ellis and McGuire (1986) showed in such a case that providers who value their own profits more than benefits to patients may underserve patients. Underservice was not a problem according to a recent RAND Corporation study of the quality of care under PPS. We found few noticeable effects of PPS on mortality and
quality of care for hospitalized Medicare patients [Rogers et al., to appear]. However, the post-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. 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.

Goodall and Pope 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 justifiable factors are too expensive to collect, and using the hospital blend is optimal under this presumption. 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.

After discussing the difficulties in separating justified from unjustified costs, I will present some estimated 'optimal' (in the Pope–Goodall sense) weights on hospital specific costs under alternate definitions of justifiable. The estimates are based on a recent study of the causes of expensive cases [Keeler (1990)].

2. What currently uncompensated cost factors are justifiable?

Certain hospitals may have Medicare patients that are inherently more expensive to treat properly because of their higher sickness (within their DRG) and other factors. We have data that allow us to show below that sickness alone is responsible for a substantial proportion of costs, but other characteristics of patients: their language, their in-hospital behavior or their lack of suitable post-hospital care might also add to costs and ideally would also be compensated.

Costs due to sickness measured at admission are clearly justifiable, as are later sickness and even bad outcomes that were not preventable. Most of these factors raise costs and so direct payment adjustments based on them might reward bad care and be susceptible to fudging in recording. However, we can use them here to compute the share of justifiable costs in hospital cost variation.

Cost savings from quick deaths should also be taken into account. Most quick deaths are not preventable [the physicians in Dubois et al. (1987), estimated that 75–85% of in-hospital deaths in that sample were not preventable]. Since quick death leads to a reduction in costs, it can even be
used in DRG payments, for example, DRG 123 (AMI with death) pays less than DRG 121 (AMI with live discharge).

3. A study of hospital cost differences

3.1. 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 are hospitalized prior to the implementation of PPS with patients hospitalized after PPS was introduced (July 1, 1985–June 30, 1986) [Draper et al., to appear]. The analysis of costs was based only on patients from the post-PPS period.

We gathered data on a nationally representative sample of 7,156 patients from the post-PPS period hospitalized with one of five diseases: congestive heart failure (CHF), acute myocardial infarction (AMI), pneumonia (PNE), cerebrovascular accident (CVA), and hip fracture (HIP). Patients with these diseases represented 19% of Medicare discharges and 32% of Medicare deaths within 30 days of admission in fiscal 1986. Patients were included if they were 65 years of age or older, had an ICD-9-CM code consistent with one of the study diseases, and their medical record confirmed that the study disease was the reason for hospital admission. The patients came from one of 296 selected acute-care general hospitals from 30 cities or towns in five states, each from a different region of the United States. The hospitals were selected to be representative with respect to urbanicity, percentage of Medicare patients, hospital size, teaching intensity, and type of ownership.

3.2. Measuring mortality and charges

By using the patient's last name, first name, date of birth, and Health Insurance Claim (HIC) 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). The HIMF contains data on dates of death subsequent to discharge. The unmatched patients were sicker on average than were matched patients, and 19% had died in-hospital.

To study charges we used the 93% of patients we had matched to the Health Insurance Master File and matched them to HCFA's Bill Retrieval File of hospitalizations (see table 1). Of the 93% of post-PPS patients for whom we had accurate mortality data, we were able to accurately match 94% to hospitalizations with nonzero charges in the Bill Retrieval File (six cases with charges = 0 in the HCFA files were dropped) for an overall matching success rate of 87%. The resulting file allowed us to measure
charges and outlier payments, and find the assigned DRG for the admissions we had abstracted.

The distribution of total charges was very skewed, so for individual level analyses, we used the log of charges. Charges were transformed to costs by adding the log of the 1984 cost to charge ratio. In the 13/296 hospitals for which the cost to charge ratio was missing, we used 0.66, the 1984 national average.

The adjusted costs per case are defined as the costs divided by the payment factors: the FY 1987 wage index (actually 0.2085 + 0.7195 \times FY87 wage index), the 1986 DRG weights, and the disproportionate share and teaching adjustments = \{dsh + 1.5 \times [(1 + irb)^{0.405} - 1]\}, where dsh is the 1987 disproportionate share factor, and irb is the 1984 interns and residents to beds ratio. Because the five diseases in the sample are serious, the average DRG weight was 1.5 and the adjusted costs have a mean of $3,130, as opposed to a mean of $4,600 for unadjusted costs.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>CHF</th>
<th>AMI</th>
<th>DNE</th>
<th>CVA</th>
<th>HIP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical RAND sample</td>
<td>1.465</td>
<td>1.437</td>
<td>1.408</td>
<td>1.442</td>
<td>1.404</td>
<td>7,156</td>
</tr>
<tr>
<td>Match to HIMF</td>
<td>1.367</td>
<td>1.329</td>
<td>1.321</td>
<td>1.342</td>
<td>1.295</td>
<td>6,665</td>
</tr>
<tr>
<td>Match to bills with nonzero charges</td>
<td>1.279</td>
<td>1.237</td>
<td>1.227</td>
<td>1.250</td>
<td>1.228</td>
<td>6,221</td>
</tr>
</tbody>
</table>

3.3. Measuring sickness at admission

To take account of 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. Variables were included if they accurately describe 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.

We used disease-specific abstraction forms to guide the collection of data about sickness at admission from the medical record of hospitalized sampled patients. 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 variables. In all instances, we collected the first available data for the patient after admission. 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 record. To take account of patients' sickness at admission, we collected 60–80 disease-specific variables from the medical record that
described the patients' acute and chronic morbid and comorbid conditions. Variables were included if they accurately describe 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.

3.4. Developing the cost-sickness scales for each disease

The scaling of important continuous and categorical variables was previously done in work to find sickness scales for each disease to predict 30-day post admission mortality [Keeler et al., to appear]. These sickness-mortality scales were used with other variables to predict costs.

We used modified stepwise backwards regression to reduce the number of sickness-at-admission variables that would be used to predict costs. The log of adjusted costs was regressed on the sickness mortality scale and the other 60-80 variables. Insignificant variables and a few significant variables with an implausible sign were dropped leaving the sickness mortality scale plus 5-12 additional variables in the scales, depending on the disease. Because the scales are the linear predictors of the log of adjusted costs, a 0.01 increase in the cost-sickness scale is associated with a 1% increase in adjusted costs, on average. At the individual patient level, the scales predicted 11-15% of the variance in adjusted log costs, depending on disease. We tested the degree of overfitting by a cross-validation technique, concluding that the conventional R-squared was a 1-2 percentage point overestimate [Keeler (1990)].

3.5. In-hospital adverse outcomes

The in-hospital adverse outcomes scale combines three parts: major complications, general sickness at day 3 (a scale of 7 vital signs and symptoms), and new lab abnormalities. Adverse outcomes that occurred on the day of death were not included, but adverse outcomes that were followed by later in-hospital death were included. Thus, it would be more precise to call them 'not immediately fatal' adverse outcomes. The overall nonfatal complications scale for each disease 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.

4. Explaining differences in average hospital costs

I presume with Pope that Medicare wants to compensate all hospitals at the same efficient level of treatment intensity, and not pay more for inefficiency or for uniformly more intense treatment at one hospital than another. Hence, I would have liked to regress adjusted costs on the justifiable
variables, while controlling for inefficiency and treatment intensity. Unfortunately, good proxies for inefficiency and intensity that are not linked to sickness were not available.

For a hospital level analysis, the adjusted costs and patient variables from the approximately 23 cases per hospital were averaged to make adjusted short run average costs and average patient variables for each hospital. These costs were regressed on payment variables: the wage index, DRG weights, disproportionate share, and teaching adjustments; and on averaged sickness, various measures of quick death, and nonfatal in-hospital complications.

Two of the payment variables, the DRG weight and the wage index were significantly related to average adjusted costs. However, this result apparently comes from their correlation with sickness, because when sickness is included, they are not significantly related ($t < 0.5$). The other two payment variables have small effects whether or not sickness is included in the regression. Thus, the current payment adjustments appear to be adequately connected to costs. Table 2 gives the $R$-squared for average adjusted costs regressed on various combinations of average patient variables (i.e., different definitions of justifiable costs). 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% of the variance in average adjusted hospital costs is explained by justifiable factors. Hence, according to the equity arguments of Pope and Goodall, 15–29% is a lower bound on the desirable weight on hospital-specific costs in a blend. Some of the remaining variance is noise, some comes from differences in treatment intensity and efficiency, and some from justifiable factors that we did not measure in the study. I tried many other hospital-level predictor variables (state, volume, urbanicity, control, percent Medicare), but only an indicator for urban government hospitals was strongly related to (lower) adjusted costs.
A blend may be more equitable, but it would put less pressure on cost control at hospitals than the current system (even if historic costs are used for the hospital-specific portion of payments, if they are ever updated the hospitals will get a portion of current expenditures back in the future). 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, it will be important to give providers some financial incentives for supplying expensive care when it is needed.

References


Keeler, E.B., 1990, Determinants of expensive cases, RAND N-3205-HCTA.


