Hospital Characteristics and Quality of Care

Emmett B. Keeler, PhD; Lisa V. Rubenstein, MD, MSPH; Katherine L. Kahn, MD; David Draper, PhD; Ellen R. Harrison, MS; Michael J. McGinty, PhD; William H. Rogers, PhD; Robert H. Brook, MD, ScD

Objective.—To compare quality of care measured by explicit criteria, implicit review, and sickness-adjusted outcomes at different types of hospitals.

Design.—Further analysis of data retrospectively abstracted from medical records to evaluate the effects of prospective payment on quality of care for hospitalized Medicare patients.

Setting.—Hospitals in five states were sampled to represent the national Medicare admissions along many dimensions.

Patients.—A total of 14 008 elderly patients with one of the following five diseases: congestive heart failure, acute myocardial infarction, pneumonia, stroke, or hip fracture. These patients were randomly sampled from those with these diseases in 297 hospitals in two time periods, 1981 to 1982 and 1985 to 1986.

Outcome Measures.—Explicit criteria, implicit review, and mortality within 30 days of admission adjusted for sickness at admission.

Results.—Quality of care ratings for hospital types are similar using explicit criteria, implicit review, and outcomes adjusted for sickness at admission. Quality differences between types of hospitals were large, with the lowest group estimated to have four percentage points higher mortality than major teaching hospitals in a cohort of patients with average mortality of 16%. Quality varies from state to state, but teaching, larger, and more urban hospitals have better quality in general than nonteaching, small, and rural hospitals. Hospital quality persists over time, but small nonteaching hospitals narrowed the gap with better quality hospitals between 1981 and 1986.

Conclusions.—The different measures led to consistent and plausible relationships between quality and hospital characteristics. Thus, valid information about hospital quality can be obtained. We need to develop ways to use such information to improve care.

(*JAMA*. 1992;268:1709-1714)

IN 1855, Florence Nightingale tabulated the death rates from leg amputation of English soldiers who participated in the Crimean War. She observed that after accounting for the level of amputation above or below the knee, soldiers operated on in large hospitals were more likely to die than those operated on in small hospitals. She identified the causes of this unexpected finding as poor sanitation and the rapid spread of infection from patient to patient in large hospitals. She pleaded with English royalty to do something about the sanitary conditions of English field hospitals.[1]

More than a century later, there is still great interest in characterizing hospitals that provide better or worse care and in using that information in some way to improve care. Many studies have shown that better outcomes for specific procedures are related to the number of such procedures that hospitals and physicians perform.[2,3] The Health Care Financing Administration (HCFA) administrative data have been used to see whether states with stringent review procedures or with more intense competition have higher in-hospital mortality rates for Medicare patients.[4] Since 1987, the HCFA has published statistics on the annual adjusted 30-day postadmission mortality rates, adjusted for case mix, for all hospitals that serve Medicare patients. These rates have also been analyzed for their association with hospital characteristics, showing that nonprofit hospitals and teaching hospitals have lower adjusted mortality than average.[5] However, the validity of comparisons made with administrative data has been questioned by researchers, clinicians, and policymakers.

In our evaluation of the effects of the diagnosis related group–based prospective payment system (PPS) on hospital quality of care,[6] we used a clinically detailed assessment of the process of care (what physicians and nurses do to patients) in addition to adjusted mortality within a fixed time period after hospital admission to assess quality of care. Data to assess process of care and sickness at admission came from a review of the patient's medical record. Here we use these data to assess how quality of care varies by hospital characteristics. In addition, we use data collected about patients admitted in 1981 to 1982 to see how varia-

From the Health Program of RAND, Santa Monica, Calif (Drs Keeler, Rubenstein, Kahn, Rogers, and Brook and Ms Harrison); the Departments of Medicine (Drs Kahn and Brook), Health Services (Dr Brook), and Mathematics (Dr Draper), University of California at Los Angeles; the Sepulveda (Calif) Veterans Affairs Hospital (Dr Rubenstein); and the California Association of Public Hospitals, Berkeley (Dr McGinty).

Reprints not available.

tions in hospital quality changed over time.

METHODS

Sample and Data From the PPS Quality of Care Study

The sample included 14 008 patients hospitalized with one of five diseases: congestive heart failure, acute myocardial infarction, pneumonia, cerebrovascular disease, and hip fracture. Half of the study patients were hospitalized in calendar year 1981 or 1982, and half were hospitalized between July 1, 1985, and June 30, 1986. The sample consists of elderly (aged 65 years or older) Medicare patients from 297 hospitals in 30 geographic areas in five states, each from a different region of the United States. Hospitals were selected to represent the national patient cohort with respect to city size, percentage of Medicare patients, hospital size, teaching intensity, patient mortality for the five diseases, and type of ownership. We oversampled hospitals caring for poor patients in order to address questions about care for the poor. Estimates weighted to reflect that oversampling differ little from the unweighted estimates presented here. In each hospital, we selected a random set of five patients for each disease admitted in each time period except for some small hospitals where fewer records were available. Demographic characteristics of the study sample and details of the sampling are described elsewhere.[7]

Measuring Mortality, Sickness at Admission, and Excess Deaths

We used already existing HCFA files to determine whether patients died in the 30 days following admission.[8] To take account of how sick patients were at admission, we collected 60 to 80 disease-specific variables from the medical record that described the patient's acute and chronic, morbid and comorbid conditions.[9-13] Based on these data we used clinical judgments and logistic regression to scale, select, and weight the most important variables making disease-specific sickness-at-admission scales that predict mortality within 30 days of admission.[14] For each patient, we used logistic regression of mortality on the scales in the post-PPS sample to compute predicted mortality, the probability that patients will die within 30 days based on their sickness at admission if they receive care at the average level for hospitals in the study for the post-PPS period. To measure excess mortality for each disease in each hospital, we subtracted the average predicted mortality from the average observed mortality. Throughout the article, reported differences in excess mortality are given in percentage points so that if a hospital has 12% observed mortality and predicted mortality is 10%, that hospital will be said to have 2% excess mortality.

The amount of information recorded in patients' medical records may influence our assessment of how sick they are. For instance, some hospitals may perform more laboratory tests and thereby discover more laboratory abnormalities than others. In addition, hospitals may vary in the detail with which patients' comorbidities are recorded by their physicians, causing patients at "high recording" hospitals to appear sicker than do similar patients in hospitals with less detailed medical records. To assess the sensitivity of results to recording practices, we also computed predicted mortality based only on first vital signs, sex, and age, which are less influenced by recording.

Measuring Process by Explicit Criteria

We initially created process measures for each disease based on literature reviews and expert opinions. Next, disease-specific panels of physicians reviewed the criteria for clinical validity and helped us select criteria that would not be vulnerable to variations in recording by year, state, or hospital type.

We developed and pilot-tested disease-specific abstraction forms to collect the medical record data. The instruments for measuring process are publicly available.[9-13]

Using clinical judgment, we grouped process criteria into scales and then tested our groupings by Likert scaling methods. These methods produced five process scales: physician cognitive diagnostic, nurse cognitive diagnostic, technical diagnostic, technical therapeutic, and monitoring with the intensive care unit or telemetry. We combined these five scales to create the overall explicit process scale reported here.

Our final process scores have been shown to be reliable and valid for measuring quality of care. Higher in-hospital process scores were associated with lower postadmission mortality, after adjusting for sickness at admission.[15]

Measuring Process by Implicit Review

All included medical records underwent explicit review, as previously described. A stratified random sample of 10% of the included medical records were selected to undergo implicit review, of which 93% were reviewed. We randomly selected one medical record per disease per hospital, with oversampling in patients who died in the hospital. In the analyses reported in this article, data have been reweighted to achieve representativeness of our findings for all Medicare patients with one of our five diseases. The reweighted implicit review sample closely matches the explicit review sample in demographic and hospital characteristics.

To perform implicit review, five physician reviewers per disease were selected by the five state professional review organizations (PROs) participating in the study. Reviewers participated in intensive training, then used a structured review form to rate medical records.

The analyses reported in this article are based on the eight-point Implicit Overall Quality Scale. This scale is the sum of responses to two Likert scale questions. The first question was, "Considering everything you know about the patient, please rate overall quality of care." The five response categories for this question ranged from extreme, above standard, to extreme, and below standard. The second question was, "Would you send your mother to these physicians in this hospital?" The four response categories ranged from definitely yes to definitely no. The scale has been shown to be reliable and to predict both explicit process scores and outcomes of individual patients.[16] To clarify the relations between the different measures of quality, we standardized both explicit and implicit process scales to have a mean of zero and a patient standard deviation of one.

The explicit abstractors and implicit raters could not be blinded to whether the patient died in the hospital. If they were influenced to score process lower for patients who died, the link between process and outcome would be artificially high. To test for this effect, we used the fact that process scores for one patient are positively correlated to those of other patients within a hospital. We took the average process for the other patients at the hospital as a proxy for the process for each patient in a separate analysis of the links between process and mortality. After adjusting for sickness, the value of the average of the other patients' process in predicting death (as quantified by logistic regression coefficients) is 75% of the value of a patient's own process on mortality for explicit process and 48% for implicit process.

Defining Hospital Groups for Comparison

We defined hospital types based on structural characteristics including size, ownership, urban or rural setting, state, size of training programs, classic city-county hospitals, and proportion of Medicaid and Medicare patients seen. The breakpoints for hospital categories were taken from the literature and were selected prior to looking at the data. Hospitals were designated as "high Medicaid" if they admitted more Medicaid patients than 90% of other hospitals in their state. PRO physicians identified within each state the classic city-county hospitals that care for many poor and uninsured patients who may have no other place to receive care. We used the area resource file and the 1984 American Hospital Association survey to determine the characteristics of study hospitals.

Table 1.—Characteristics of the 297 Study Hospitals

	Mean	SD	Minimum	Maximum
1984 Hospital characteristics				
No. of beds	308	230	20	1569
Intern and resident/bed ratio	0.07	0.14	0	0.80
Medicaid admissions, %	12	13	0	67
Medicare admissions, %	34	12	6	80
Average of quality measures in 1985 to 1986				
Explicit process	0.0	0.40*	−1.29	0.97
Implicit process	0.0	0.70†	−2.56	2.77
Mortality within 30 d, %	15.9	7.7*	0.0	40.0
Predicted mortality, % ‡	15.9	4.3*	7.0	31.2
Excess mortality, % §	0.0	6.6	−19.2	22.5

*SD of hospital averages of our sample of approximately 22 cases per hospital in 1985 to 1986. The SDs of averages over all patients in the hospitals would be considerably smaller.

†Based on hospital averages of two to three cases per hospital in 1985 to 1986.

‡Predicted mortality is based on sickness at admission to the hospital.

§Calculated by subtracting the average predicted mortality from the average actual mortality.

Statistical Methods

Hospital average values for measures were obtained by averaging patients within a disease and by averaging the five disease measures. Preliminary analyses had shown relations between quality and hospital characteristics were similar across diseases. The unadjusted association between hospital characteristics and measures of quality were tested by F tests.[17]

To decide which categories of hospitals to report, we used a computerized recursive partitioning algorithm (as implemented in Treetools[18]) to construct hospital quality regression trees. Tree-based models have been used widely in medical studies.[19] Such models do not prespecify the form of interactions and are easy to interpret and cross-validate. The algorithm first identifies the single hospital characteristic and a division of that characteristic that most strongly differentiates across hospitals average explicit process; it then divides the hospitals into two groups based on levels of that characteristic, each group having more similar quality than the original population. Each subgroup is then divided in the same way (using the same or another characteristic), and the process continues until subgroups are too small to be split further. The algorithm then uses 10-fold cross-validation to prune the full tree back to the largest statistically stable tree. Splitting rules are developed on calibration subsamples consisting of a random 90% of hospitals and tested for predictive accuracy on the other 10% of hospitals. This procedure is carried out for all 10 distinctive 90%-10% splits, and the predictive accuracy measures are averaged. Splits with low predictive accuracy are then pruned. The validation procedure is analogous to significance testing of coefficients in regression but avoids problems of variable selection that may introduce bias. Trees are grown twice, once allowing any split of the continuous variables (beds, amount of teaching, percentages of Medicare and Medicaid admissions) and once restricting categories to our preselected cutpoints. Except for the split on amount of teaching, we used our preselected categories throughout. The proportions of Medicaid and of Medicare admissions were not related to quality, and results based on those characteristics are not reported here.

Except for results on changes in quality of care over time, all results presented here are based on the patients in the 1985 to 1986 sample for whom mortality could be checked (n=6665). The numbers of patients studied from each hospital are close to constant (mean, 22.4; SD, 3.7), so results weighted by patients per hospital are quite similar to the unweighted results presented here.

RESULTS

The distribution of hospitals by selected characteristics is shown in Table 1. Hospitals in our sample closely match national data in terms of rural vs urban location, size, teaching intensity, and proportion of Medicare and Medicaid patients.[7]

Comparisons Across Measures of Quality of Care

The unadjusted mean values of the major measures of quality for subsets of hospitals are shown in Table 2. Overall, each measure of quality has a mean of zero, so that a positive value on explicitly or implicitly measured process indicates above average care, as does a negative value for excess deaths. There is remarkable agreement among the three measures of quality. For all 19 listed hospital characteristics except for-profit status and bed sizes of 100 to 400, above-average explicit process (ie, a positive sign) is associated with above-average implicit measured process and lower mortality than expected, and below-average process with higher mortality than expected. In the exceptional cases, explicit process is very close to average. Because of the agreement, we can think of any of the three measures as representing "quality."

Implicit process differs more than our explicit process measures by hospital type. According to the implicit process measures, the average patient at the major teaching hospital receives care in the 75th percentile of patients overall (0.66 is the 75th percentile of a standard normal deviate), but in the 61st percentile of patients overall as measured by explicit process (0.28 is the 61st percentile of a standard normal deviate).

Comparisons Across Types of Hospitals

Rural hospitals provided care that was below average. To be precise, in the 61 rural hospitals, mean explicit process was −0.30, mean implicit process was −0.33, and mean excess death was +1.4%. This means the average patient sampled from a rural hospital was given care that was 0.30 SD worse than that given the overall average patient as measured by explicit process criteria and was 0.33 SD worse than average as measured by implicit process criteria. These patients had a 1.4% greater chance of dying than if they had been treated in the average hospital. The excess mortality of 1.4% is the difference between observed mortality of 17.4% and predicted mortality based on sickness at admission of 16.0%. Our data contain only a few cases from each hospital, but the many hospitals in each subset make each clinically significant difference in quality statistically significant. The city size, teaching, state, and hospital size classifications lead to differences in explicit and implicit process that are significant at the .001 level and to differences in excess death that are significant at the .05 level.

Table 2 shows that hospitals in bigger

Table 2.—Average Quality at Hospitals With Various Characteristics

Hospital Characteristics	Distribution of Hospitals, % *	Explicit Hospital Process, SD	Implicit Hospital Process, SD	Excess Mortality, %	Observed Mortality, %	Predicted Mortality, %
City size						
Rural	21	−0.30‡	−0.33‡	1.4§	17.4	16.0
SMSA <1 000 000†	33	−0.03	−0.11	0.6	16.1	15.5
SMSA >1 000 000	46	0.16	0.22	−1.1	15.1	16.3
Teaching status						
Major teaching‖	10	0.28‡	0.66‡	−2.5§¶	14.7	17.2
Moderate and limited	26	0.13	0.21	−0.8	14.9	15.7
Nonteaching	64	−0.09	−0.18	0.7	16.5	15.9
State						
A	20	−0.23‡	−0.25‡	2.5§	17.6	15.0
B	20	−0.18	−0.14	−0.0	16.7	16.8
C	20	0.02	−0.00	−0.4	15.1	15.5
D	20	0.15	0.10	−1.3	14.0	15.3
E	20	0.23	0.38	−0.9	16.3	17.1
Ownership						
Nonprofit	62	0.06‡	0.07§	−0.5	15.3	15.8
For-profit	12	0.04	−0.15	−0.4	16.0	16.5
Public	26	−0.17	−0.12	1.2	17.4	16.2
Size, No. of beds						
<100	19	−0.30‡	−0.40‡	1.9§	17.9	15.9
100-200	21	−0.01	−0.09	−0.5	15.6	16.1
201-400	31	0.07	−0.04	−0.1	15.3	15.3
>400	30	0.12	0.32	−0.8	15.6	16.5
Classic city-county#	13	0.13§	0.28§	−1.3	15.7	17.0
Total	**100**	...	...	...	**15.9**	**15.9**

*To estimate the number of patients measured in each group, multiply the percentage in the group by 6665 patients for explicit process and excess mortality and by 654 patients for implicit process.
†SMSA indicates standard metropolitan statistical area.
‡Classification leads to differences that are significant at the .001 level.
§Classification leads to differences that are significant at the .05 level.
‖Major teaching is defined by a ratio of interns and residents to beds greater than 0.27 in 1984.
¶Using predicted mortality based only on vital signs, sex, and age (to avoid a possible recording bias), this figure would be −1.1%.
#Identified by professional review organization physicians in each state.

cities and certain states have better average quality. Better quality is also associated with more teaching, private ownership, and bigger hospitals. The big, classic city-county hospitals that serve the urban poor do well on average, perhaps because most of them in our sample are teaching hospitals. Nonprofit and for-profit hospitals provide similar quality overall.

The major teaching hospitals have sicker patients at the time of hospital admission (ie, higher predicted mortality) on average. Major teaching hospitals had many more patients with laboratory abnormalities and comorbidity, but a similar proportion of patients with poor vital signs as did the other hospitals. We cannot determine from our data the extent to which teaching hospitals have truly sicker patients as opposed to more laboratory ordering and recording of comorbid conditions, but a recent study of 15 Boston hospitals found severity to be somewhat higher for patients in teaching hospitals.[20] According to the sickness scales based only on vital signs, age, and sex, the patients at teaching hospitals have sickness at admission similar to that of patients admitted to other hospitals and a 1.1% lower adjusted mortality rate than did the average hospital.

We next show the combined impacts of these characteristics on quality. The regression tree shows large differences in quality due to teaching status and amount of teaching, and for nonteaching hospitals, differences across states, big city location, public vs private, and bed size (Table 3). At each split, the higher branch is associated with better quality. As in Table 2, there is striking consistency of the three measures of quality, with poorer explicit and implicit process associated with excess mortality.

The teaching hospitals do not split on the other factors, reflecting the fact that teaching hospitals are less influenced by location and ownership. Indeed, the variance of measured hospital average explicit process in teaching hospitals is half that in nonteaching hospitals. The split within teaching hospitals occurs at only 0.062 intern and resident per bed, with hospitals doing a limited amount of teaching having average quality and the hospitals doing more teaching having considerably better quality.

The nonteaching hospitals vary greatly from state to state. However, regardless of state, quality is worse outside big cities, in public hospitals, and in hospitals with less than 100 beds. In the 31 hospitals at the bottom of the tree, average explicit process was −0.54, implicit process was −0.47 and excess mortality was 2.6%. The difference in excess mortality between these hospitals and the major teaching hospitals is 4.3% (2.6%+1.7%). If we used only vital signs, age, and sex to measure sickness, this difference would not change.

Changes Over Time

The changes in quality that occurred between the two time periods, 1981 to 1982 and 1985 to 1986, are shown in Table 4. To avoid confounding from changes in recording and laboratory ordering, predicted mortality was based only on vital signs, sex, and age. The marked improvements in the total for all three measures of quality have been shown before.[8,15,16]

The regression tree has only three branches because few of these categorical variables affect changes in quality strongly. Measurements of changes in quality are less precise than measurements of 1985 to 1986 hospital quality because they are affected by random variation in each period, and because quality was more variable in 1981 to 1982. Each of the three hospital groups shown improved markedly. The groups

Table 3.—Average Quality in Categories Defined by a Regression Tree of Explicit Progress on Hospital Characteristics*

Hospital Categories	No.	Explicit Process	Implicit Process	Excess Mortality
Major and moderate teaching	69	0.28	0.51	−1.7
Limited teaching (<0.062 intern and resident/bed)	37	−0.04	0.03	−0.4
Nonteaching				
States C, D, E				
Big city	45	0.15	−0.03	−0.7
Not big city				
Private	41	0.11	0.01	−0.0
Public	16	−0.18	−0.30	1.7
States A, B				
Big city	20	0.02	−0.25	−1.4
Not big city				
≥100 beds	38	−0.27	−0.32	2.2
<100 beds	31	−0.54	−0.47	2.6
Total for All Hospitals	**297**	. . .	. . .	. . .

*Categories were defined by a regression tree of average explicit process at hospitals with potential categories given by number of beds, amount of teaching, percentage of Medicare and Medicaid admissions, city-county status, state, standard metropolitan statistical area (SMSA) size, and ownership. Big city is defined as in an SMSA with greater than 1 000 000 population; private indicates nongovernment hospitals.

Table 4.—Differences in Average Hospital Quality Between 1981 to 1982 and 1985 to 1986 in Categories Defined by a Regression Tree*

Hospital Categories†	No.	Explicit Process	Implicit Process	Excess Mortality
Any teaching	106	0.30	0.36	−1.5
Nonteaching				
≥200 beds	85	0.36	0.39	0.9
<200 beds	106	0.54	0.36	−0.7
Total for All Hospitals	**297**	**0.41**	**0.36**	**−0.5**

*Entries represent average 1985 to 1986 values minus 1981 to 1982 values of each measure, averaged over the hospitals in the category.

†Categories were defined by a regression tree of changes in average explicit process at hospitals with potential categories given by number of beds, amount of teaching, state, standard metropolitan statistical area size, and ownership.

that were behind caught up somewhat in explicit process, but not in implicit process or excess death. The range of the changes over time shown in Table 4 and in the unpruned tree we did not show is much smaller than the differences among hospital types shown in Table 3, reflecting the stability of hospital quality over time. The patterns of quality in 1985 to 1986 were also seen in the 1981 to 1982 data, but quality differences were even larger in the earlier period.

COMMENT

Consistency of Different Measures of Quality

The most striking result from this analysis is the consistency of the three clinically based measures of quality across hospital categories. Earlier studies that have had difficulty in showing relationships between process and outcomes were hindered by the lack of clinical detail in, for example, HCFA's administrative data [21] or by less developed measures of process and sickness.[22] We were fortunate to have the resources to collect the large amounts of clinically detailed data on process and sickness that are necessary to establish these links.

The consistency of the measures at the hospital level reinforces our earlier findings of strong relationships among quality measures at the patient level.[14-16] Moreover, the strong relationship of sickness-adjusted mortality for each patient to average process of other patients at the hospital is evidence that the process-outcome link is not due to bias in assessing quality for patients who died. In these data, logically defined hospital groups with better explicit process had better implicit process and lower mortality after adjustment for sickness at admission. These relationships represent real effects rather than statistical or methods artifacts.

At the individual hospital level, our samples were small. About 22 patients had their medical records abstracted to determine explicit process and sickness at admission, and only two to three patients had the structured implicit review to determine the results in Table 2. Even if there were no differences in hospital quality and patient survival was random with constant death probability and predicted death of 16%, observed mortality of groups of 22 patients would vary around 16% with an SD of 7.8% so that one sixth of all hospitals would have excess mortality greater than 7.8%. Our sample is too small to say convincingly that a particular hospital is good or bad. However, aggregating our data over, for example, the 30 major teaching hospitals in the study would bring the SD of observed mortality under this scenario down to 1.4%, which is precise enough to show the teaching hospitals to be superior. The agreement seen throughout Tables 2 and 3 between the two clinical measures of process and the outcome measure of 30-day mortality adjusted for sickness at admission is evidence for the validity of all three measures. It allows us to talk simply about results on quality of care in different groups of hospitals without specifying whether we mean process or outcome.

Variation in Hospital Quality

Quality differences between the best and worst groups of hospitals were dramatic. The estimated expected increase in mortality for study patients of a given sickness who go to the lowest quality group of hospitals instead of the major teaching hospitals is over four percentage points. Quality differences among hospitals within the groups increase these discrepancies. Considering that the overall 30-day postadmission mortality rate was 15.9%, these differences are substantial enough to cause us to think seriously about how care in the less favored hospitals and hospital groups can be improved.

Quality improved steadily with the size of hospital and the population of the community in which it was located. To see which aspects of care were substandard we computed the rural hospital averages for process for each disease and for the five subscales of the explicit process measure (data not shown). Process averages were similar across diseases, but there were big differences across the subscales. At rural hospitals, the physician cognitive diagnostic score was the worst subscale, technical diagnostic and monitoring scores were intermediate, and nurse cognitive and technical therapeutic scores were about the same as in hospitals overall. Rural hospitals may have difficulty in attracting skilled physicians, and these physicians may not have enough patients or contact with other physicians to maintain their skills.

Despite their generally lower quality, rural hospitals can symbolize a small town's identity, provide the means to attract physicians that will make medical care more convenient, and provide jobs for the community. Thus, small towns will try to keep hospitals from closing even if many people in the community will not use them when they are seriously sick. However, the Department of Health and Human Services Inspector General recently noted that 75% of the people affected by the 50 rural hospitals that closed in 1988 were within 32 km of another hospital.[23] In a review of this dilemma, Ermann[24] suggested that small-town hospitals might diversify into home health, nursing home, and ambulatory services and affiliate with urban institutions. Some underoccupied, low-quality hospitals might be closed.

The quality of care was strongly related to teaching in our data, with more teaching associated with better quality. This may be due in part to teaching hospitals being generally large and located in bigger cities. Alternatively, it may have to do with more intense continuing medical education for physicians and nurses at teaching hospitals. In contrast to the stories of chaos we sometimes hear from the media, public teaching hospitals in 1986 had better process than private hospitals, and the city-county hospitals had generally high quality, perhaps because most were large teaching hospitals.

Our results generally concur with those of Hartz et al,[5] who also found that mortality was lower at more urban hospitals and at hospitals with more teaching activity, and higher in public hospitals. In contrast to Hartz et al, we found that quality at for-profit hospitals was similar to quality at private nonprofit hospitals. The slightly better quality at for-profit nonteaching hospitals makes up for the much higher percentage of (high-quality) nonprofit hospitals with teaching. Shortell and Hughes[4] in their study of in-hospital mortality at 981 hospitals in 1983 to 1984 also found no differences by ownership.

There was general improvement in quality between 1981 to 1982 and 1985 to 1986, but results on which hospitals improved the most are mixed. The lower quality hospitals narrowed the advantage of the teaching hospitals in explicit process but not in excess mortality. To avoid biasing comparisons of quality before and after the introduction of the diagnosis related group PPS and among different types of hospitals, our measures of explicit process were chosen to reflect care that should have been offered through the 5 years of the study and in all types of hospitals. These basic measures include, for example, documentation of thorough history and physical for each patient. Perhaps the small rural hospitals improved more in explicit criteria than in outcomes because the better documentation observed in 1985 to 1986 may not be as important to outcomes in small hospitals as in large hospitals. One might also speculate that teaching hospitals do better in 1985 to 1986 in mortality and in implicit process than might be expected from our measures of explicit process because they have adopted new advances in technology more quickly than other hospitals. This is possible, but our data show use of such therapies as streptokinase in patients with a myocardial infarction was low even in teaching hospitals in the later period (5.5% overall).

The data analyzed herein are limited to selected measures of quality on a sample of patients at a sample of hospitals. The patients were elderly Medicare patients with one of five serious diseases. Pediatric, obstetric, and psychiatric services were unrepresented, and the only surgical condition was hip fracture. There were considerable differences in quality by states, so that our sample of five states is insufficient to estimate national performance on these measures. The measures of process were selected among other reasons for their probable association with mortality. For other diseases such as cataracts or for terminal care, mortality would not be a primary outcome of interest, and other process measures should be chosen.

The consistency of evaluation by the different measures of quality and the systematic plausible relationships of quality to hospital types support the idea that valid information about hospital quality can be obtained. The challenge is to develop ways to use that information to improve care for hospitals across the spectrum of quality.

This work was supported by cooperative agreement 18-C-98853/9-03 issued by the Health Care Financing Administration, Baltimore, Md, and by grant 1 R01-HSO-6546 from the Agency for Health Care Policy and Research, Rockville, Md. We thank Robert Bell, PhD, and Sally Morton, PhD, for help with statistical issues, and Katherine Desmond for meticulous statistical programming.

References

1. Nightingale F. *Notes on Nursing: Enlarged and for the Most Part Rewritten.* 3rd ed. London, England: Longman, Green, Longman, Roberts and Green; 1863.
2. Luft HF, Garnick DW, Mark DH, McPhee SJ. *Hospital Volume, Physician Volume, and Patient Outcomes.* Ann Arbor, Mich: Health Administration Press Perspectives; 1990.
3. Kelly JV. What do mortality studies reveal about hospital volume, teaching status and ownership? *Int J Technol Assess Health Care.* 1990;6:239-252.
4. Shortell SM, Hughes EFX. The effects of regulation, competition, and ownership on mortality rates among hospital inpatients. *N Engl J Med.* 1988;318:1100-1107.
5. Hartz AJ, Krakauer H, Kuhn EM, et al. Hospital characteristics and mortality rates. *N Engl J Med.* 1989;321:1720-1725.
6. Kahn KL, Rubenstein LV, Draper D, et al. The effects of the DRG-based prospective payment system on quality of care for hospitalized Medicare patients: an introduction to the series. *JAMA.* 1990; 264:1953-1955.
7. Draper D, Kahn KL, Reinisch EJ, et al. Studying the effects of the DRG-based prospective payment system on quality of care: design, sampling, and fieldwork. *JAMA.* 1990;264:1956-1961.
8. Kahn KL, Keeler EB, Sherwood MJ, et al. Comparing outcomes of care before and after implementation of the DRG-based prospective payment system. *JAMA.* 1990;264:1984-1988.
9. Kahn KL, Chassin MR, Rubenstein LV, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Congestive Heart Failure.* Santa Monica, Calif: RAND; 1988. Publication N-2798-HCFA.
10. Kosecoff J, Rubenstein LV, Kahn KL, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Acute Myocardial Infarction.* Santa Monica, Calif: RAND; 1988. Publication N-2799-HCFA.
11. Roth CP, Kahn KL, Sherwood MJ, Rubenstein LV, Chew CF, Brook RH. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Pneumonia.* Santa Monica, Calif: RAND; 1988. Publication N-2801-HCFA.
12. Rubenstein LV, Kosecoff J, Kahn KL, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Cerebrovascular Accident.* Santa Monica, Calif: RAND; 1988. Publication N-2802-HCFA.
13. Sherwood MJ, Kahn KL, Kosecoff J, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Hip Fracture.* Santa Monica, Calif: RAND; 1988. Publication N-2800-HCFA.
14. Keeler EB, Kahn KL, Draper D, et al. Changes in sickness at admission following the introduction of the prospective payment system. *JAMA.* 1990; 264:1962-1968.
15. Kahn KL, Rogers WH, Rubenstein LV, et al. Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospective payment system. *JAMA.* 1990; 264:1969-1973.
16. Rubenstein LV, Kahn KL, Reinisch EJ, et al. Changes in quality of care for five diseases measured by implicit review, 1981 and 1986. *JAMA.* 1990;264:1974-1979.
17. Scheffe H. *The Analysis of Variance.* New York, NY: John Wiley & Sons Inc; 1959.
18. Clark LA, Pregibon D. Tree-based models. In: Chambers JM, Hastie TJ, eds. *Statistical Models in S.* Pacific Grove, Calif: Wadsworth & Brooks/Cole Advanced Books & Software; 1991.
19. Goldman L, Cook EF, Brand DA, et al. A computer protocol to predict myocardial infarction in emergency department patients with chest pain. *N Engl J Med.* 1988;318:797-803.
20. Iezzoni LI, Shwartz M, Moskowitz MA, Ash AS, Sawitz E, Burnside S. Illness severity and costs of admissions at teaching and nonteaching hospitals. *JAMA.* 1990;264:1426-1431.
21. Park RE, Brook RH, Kosecoff J, et al. Explaining variations in hospital death rates: randomness, severity of illness, quality of care. *JAMA.* 1990; 264:484-490.
22. Richards T, Lurie N, Rogers WH, Brook RH. Measuring differences between teaching and nonteaching hospitals. *Med Care.* 1988;26(suppl): 1-141.
23. *Med Health.* May 14, 1990; 44(19):2.
24. Ermann DA. Rural health care: the future of the hospital. *Med Care Rev.* 1990;47:33-73.