A RAND NOTE

A Summary of the Effects of DRG-Based Prospective Payment System on Quality of Care for Hospitalized Medicare Patients

Katherine L. Kahn, David Draper,
Emmett B. Keeler, William H. Rogers,
Lisa V. Rubenstein, Jacqueline Kosecoff,
Marjorie J. Sherwood, Ellen J. Reinisch,
Maureen F. Carney, Caren J. Kamberg,
Stanley S. Bentow, Kenneth B. Wells,
Harry Savitt, Harris Allen, David Reboussin,
Carol P. Roth, Carole Chew, Robert H. Brook

October 1990

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. 18-C-98853/9-03.

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

A RAND NOTE

N-3132-HCFA

A Summary of the Effects of DRG-Based Prospective Payment System on Quality of Care for Hospitalized Medicare Patients

Katherine L. Kahn, David Draper,
Emmett B. Keeler, William H. Rogers,
Lisa V. Rubenstein, Jacqueline Kosecoff,
Marjorie J. Sherwood, Ellen J. Reinisch,
Maureen F. Carney, Caren J. Kamberg,
Stanley S. Bentow, Kenneth B. Wells,
Harry Savitt, Harris Allen, David Reboussin,
Carol P. Roth, Carole Chew, Robert H. Brook

October 1990

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

RAND

The Effects of the DRG-Based Prospective Payment System on Quality of Care for Hospitalized Medicare Patients

An Introduction to the Series

Katherine L. Kahn, MD; Lisa V. Rubenstein, MD, MSPH; David Draper, PhD; Jacqueline Kosecoff, PhD; William H. Rogers, PhD; Emmett B. Keeler, PhD; Robert H. Brook, MD, ScD

In 1985, we began a 4-year evaluation of the effects of the diagnosis related groups–based prospective payment system on quality of care for hospitalized Medicare patients. This article provides an overview of the study's background, aims, design, and methods. We used a clinically detailed review of the medical record supplemented by data on postdischarge outcomes drawn from the files of the Health Care Financing Administration and fiscal intermediaries to (1) compare outcomes of care after adjustment for sickness at admission, (2) assess the process of in-hospital care and relationships between processes and outcomes, and (3) assess status at discharge for a nationally representative sample of patients hospitalized before and after prospective payment was implemented.

(*JAMA*. 1990;264:1953-1955)

THE HEALTH care costs of the Medicare program to the US government have risen substantially over the last 25 years. Under the retrospective payment system in effect during Medicare's first two decades, Medicare expenditures rose at a much higher rate than background inflation.[1,2]

In 1982, after review of the financial incentives associated with the retrospective payment system, the US Congress passed the Tax Equity and Fiscal Responsibility Act, which placed a prospective cap, beginning in October 1982, on annual operating revenues per inpatient Medicare case at each hospital. This was altered in October 1983 by the introduction of the present prospective payment system (PPS) in which hospitals are paid an amount based largely on flat rates per admission calculated for each of approximately 470 diagnosis related groups. At the same time, the current system of monitoring of quality and appropriateness of care by professional review organizations was established.

The new payment system has been successful at slowing the upward spiral of Medicare costs.[3] However, because prospective payment contains incentives to decrease length of stay and substitute lower-cost services and procedures, concern has arisen among patients, physicians, and policymakers that, despite the introduction of monitoring by professional review organizations, the quality of care offered to Medicare patients may have declined. Since 1985 we have been conducting a national study funded by the Health Care Financing Administration of the US Department of Health and Human Services to examine the effects of the PPS on quality of care for hospitalized Medicare patients.

We had two audiences in mind in developing this study: clinicians who would like to improve the quality of care for hospitalized patients and policymakers who wish to improve the health care system. The goals of our study were (1) to assess the quality of in-hospital care for Medicare patients aged 65 years and older prior to and subsequent to the implementation of the PPS, and (2) to estimate the effects of the PPS intervention on quality of care by comparing quality of care now with our best estimate of what it would have been in the absence of the PPS. Other articles in this series present our findings.[4-10] In this introductory article, we provide an overview of our methods, offer a review of the recent literature on the effects of the PPS, comment on the generalizability of our findings, and mention some caveats on interpreting our results.

METHODS

We were not able to conduct a prospective controlled trial of the effects of the PPS on quality of care since virtually the entire country was put on prospective payment at the same time.[4] Instead, we designed a retrospective before-after study in which we contrasted data on 16 758 Medicare patients who were hospitalized in one of five states prior to and subsequent to 1983. We selected calendar years 1981 and 1982 as our pre-PPS study period and July 1985 to June 1986 as our post-PPS period, concentrating half of our sample in each of these periods.

How should quality of care pre- and post-PPS be measured? In looking for differences in quality of care across time periods or hospitals, it is natural to examine patient outcomes such as mortality. But patient sickness at admission also needs to be adjusted for in making outcome comparisons because changes in the burden of illness at admission could explain differences in outcome. Process—what clinicians do to patients—is also important. If outcomes adjusted for sickness at admission change and we do not examine the process of care, how will we know why the change occurred, and which aspects of care are now better and which are worse? Thus, we compared outcomes pre- and post-PPS after adjusting for

From the Health Program, the RAND Corp, Santa Monica, Calif (Drs Kahn, Rubenstein, Draper, Rogers, Keeler, and Brook); the Departments of Medicine (Drs Kahn, Rubenstein, Kosecoff, and Brook) and Health Services (Drs Kosecoff and Brook), UCLA; and Value Health Sciences Inc, Santa Monica, Calif (Dr Kosecoff).

The opinions, conclusions, and proposals in this article are those of the authors alone and do not necessarily represent the views of any of these organizations, the RAND Corp, or UCLA.

Reprint requests to the RAND Corp, 1700 Main St, PO Box 2138, Santa Monica, CA 90406-2138 (Dr Kahn).

sickness at admission. We measured the quality of the process of care by means of explicit criteria and by implicit review by expert clinicians. We established process-outcome links in which better process of care was shown to be associated with better outcomes. Finally, because of the PPS incentive to reduce length of stay, we also measured patient instability at discharge.

We based our evaluation of changes in quality of care on six diseases: congestive heart failure, acute myocardial infarction, pneumonia, cerebrovascular accident, hip fracture, and depression. In this series of articles,[4-10] we report on results from the first five of these diseases (sample size for the five diseases combined was 14 012); the analysis of the depression data (sample size 2746) has not yet been completed.

To identify comparable patients pre- and post-PPS, we defined each disease by *International Classification of Diseases, Ninth Revision, Clinical Modification* codes, so that patients who truly had one of the aforementioned six diseases in either period as the principal reason for admission should have been assigned one of our codes even if there were coding changes over time. We then used strict clinically detailed inclusion criteria to select patients within each disease category.

We used both explicit and implicit measures to assess the process of care. With explicit measurement, each patient's care was compared with predetermined criteria. With implicit measurement, each patient's medical record was assigned a quality of care rating based on a physician's judgment of the adequacy of the care.

After the data were collected, we used regression methods to construct sickness-at-admission scales and to adjust outcomes for sickness at admission. Clinical judgment and Likert scaling were used to construct explicit scales measuring the appropriateness of the processes of care and the level of instability at discharge. Some criteria that went into these scales were relevant to all patients and some only to subsets of patients depending on their clinical needs. We concentrated on aspects of care for which standards of good clinical practice were both unambiguous and stable over time from 1981 to 1986.

In every phase of the study we used a multidisciplinary approach. We drew on clinical expertise in general internal medicine, geriatrics, cardiology, pulmonary medicine, infectious diseases, neurology, orthopedics, and psychiatry. We enlisted physicians in each of our five study states, including physicians from both urban and rural hospitals. This clinical perspective was complemented by expertise in statistics, psychometrics, economics, health policy, and evaluation sciences.

The remaining seven articles in this series give additional details on our design, sampling, and fieldwork,[4] the measurement of sickness at admission,[5] the *explicit* measurement of process of care,[6] the *implicit* measurement of process of care,[7] patient status at discharge,[8] patient outcomes before and after the introduction of the PPS,[9] and a summary of the effects of prospective payment on quality of care.[10]

REVIEW OF THE LITERATURE ON THE PPS EFFECTS

When we began our study in 1985, little was known about the effects of the PPS on quality of care. Since then, however, other investigators have obtained results that serve as context for this series of articles.

Using a sample of 646 US nonfederal, short-term general hospitals from 1980 to 1985, DesHarnais et al[11] found that Medicare discharges and length of stay declined significantly after the introduction of the PPS, while use of skilled nursing facilities and home health care increased post-PPS. Mayer-Oakes et al,[12] studying patients in the intensive care units of three hospitals, found a 31% decrease in the number of intensive care unit beds; for patients in the intensive care unit, they found no changes in either patient severity of illness or treatment intensity after the PPS was implemented. They found a reduction in length of stay, both for the overall hospitalization and for the use of the intensive care unit. Despite the reductions in length of stay, mortality in-hospital or at 6 months after hospitalization did not change. More recently, Sager et al[13,14] used age-specific national mortality data from 1981 through 1985 to report changes in the location of death after prospective payment; place of death for some patients had moved from the acute hospital to the nursing home. In 1988, Guterman et al[3] found a reduction in the number of Medicare short-stay hospital admissions, a reduction in length of stay, and an increase in use of discharges to skilled nursing facilities using the Health Care Financing Administration's Medicare Statistical System as a data source.

Fitzgerald et al,[15,16] in their analyses of hip fracture patients in two large hospitals, reported a reduction in the number of physical therapy sessions and a reduction in length of stay. They found an increase from 38% to 60% in the proportion of hip fracture patients discharged to a nursing home, and an increase from 9% to 33% in the proportion of patients with continued nursing home care at 1 year after discharge.[16] Palmer et al,[17] however, found no change in either the proportion of hip fracture patients in one hospital discharged to a nursing home or the proportion in a nursing home at 6 months. Gerety et al[18] studied hip fracture patients in a tertiary care setting and found shorter length of stay and decreased functional status at discharge post-PPS. Their assessment of outcomes at 1 year after hospitalization showed no difference pre- and post-PPS in the fraction of institutionalized patients. These results emphasize the need for generalizable evidence about the effects of the PPS on quality of care.

GENERALIZING FROM OUR STUDY DISEASES

In choosing our six study diseases, we developed the following selection criteria: high prevalence, high mortality (so that changes in patterns of care leading to preventable death might be evident), well-defined diagnostic criteria that are readily accessible from data in the medical record, high likelihood of a strong process-outcome link (so that changes in outcome can be related to changes in process), and relative stability regarding what constitutes good care across the entire time period from 1981 to 1986.[19] We also hoped to include medical, surgical, and psychiatric conditions. Four medical conditions (congestive heart failure, acute myocardial infarction, pneumonia, and cerebrovascular accident), one surgical condition (hip fracture), and one psychiatric condition (depression) were selected. The introduction in recent years of thrombolytic agents in the care of acute myocardial infarction may have altered its outcome and thus reduced its eligibility for inclusion in our study. However, the use of these agents in 1985 through 1986 was limited and their efficacy in reducing deaths in people over 65 years of age has not been demonstrated.[20]

We can generalize our results to the population of all Medicare patients with one of our five nonpsychiatric diseases, a population that included 18% of all Medicare admissions and 32% of all deaths within 30 days of admission in fiscal year 1986.[21] But to what extent do our results generalize to Medicare patients with other diseases?

In selecting our conditions, we picked diseases for which changes in process can affect mortality. We cannot comment on changes in quality of care produced by the PPS for patients with terminal conditions such as end-stage cancer. For such patients, quality of care would need to reflect the quality of

the dying experience, and we did not measure this aspect of care.

Our results probably generalize more readily to other medical conditions than to other surgical and psychiatric diseases. In addition, insofar as we are able to show consistent changes in process of care across all of our conditions, generalization becomes more valid.

Finally, our selection of diseases makes possible a kind of upper bounding or *a fortiori* argument of the following form: by concentrating on conditions with high mortality and likely strong process-outcome links (ie, mortality can be prevented), we chose diseases for which one has reason to think that, if quality of care has declined under the PPS and mortality has increased, this decline would be likely to manifest itself in either process or outcome changes. Therefore, if we do not see a deterioration in mortality statistics, the overall detrimental effects of the PPS aggregated across Medicare patients with other diseases cannot be extremely large.

CAVEATS ON THE INTERPRETATION OF OUR RESULTS

In addition to the above remarks about generalizing from our diseases, three other major caveats are worth bearing in mind as this series of articles is read:

1. Our design involves a direct comparison between the quality of care given in 1981 through 1982 and in 1985 through 1986. We are able to associate changes in quality with the introduction of the PPS, but definitively identifying which of those changes were caused by the PPS is more difficult, since other aspects of medical care besides the PPS may also have changed from 1981 to 1986.[22] We have attempted to confront this by measuring quality of care at multiple time points in both the pre- and post-PPS periods, so that secular trends in medicine that were in place before the PPS was implemented may be at least roughly estimated and removed. Our approach in this series of articles is to present straightforward pre- and post-PPS comparisons (ie, associations) in all of the articles of this series,[4-9] except the last one, and to sort out issues of trend and causality in the last article.[10]

2. We studied quality of care only for hospitalized patients and did not examine questions of access to hospitalization prior to or following the acute hospital stay.

3. Finally, our post-PPS data are from 1985 and 1986. Although all hospitals had changed to a PPS by 1985, payments were still determined in part by the hospital's own costs through 1986. During 1985 and 1986, most hospitals were still profiting under the PPS. Since that time, PPS payments have been tightened, and yearly Medicare payment increases to hospitals are no longer keeping up with the inflation rate.[3,23,24] On the other hand, changes in length of stay associated with the introduction of the PPS appear to have stabilized and have not continued to decline in recent years.[25] Our results, therefore, may be suggestive of the state of affairs in 1990, but it seems wise to continue to collect clinically detailed data to monitor sickness at admission, processes, and outcomes on a regular basis as long as prospective payment is in place.

This series presents the results of our analyses of the impact of the diagnosis related groups–based PPS on quality of care for hospitalized Medicare patients.

This work was supported by cooperative agreement 18-C-98853/9-03 issued by the Health Care Financing Administration of the US Department of Health and Human Services, Baltimore, Md. The entire RAND costs ($3.9 million) of this project were funded with federal money.

We gratefully acknowledge the collaborative effort of the five professional review organizations that enabled this work to be completed. Participants from these organizations included medical directors, physician consultants and reviewers, project directors, and review coordinators from each of the five study states. In particular, we appreciate the keen clinical insight of the professional review organizations' physician specialists with whom we consulted throughout this study. We also acknowledge the many contributions of Harry Savitt, PhD, project officer from the Office of Research and Demonstrations of the Health Care Financing Administration of the US Department of Health and Human Services, whose administrative skills, astute commentary, and continuing support helped us immeasurably. Maureen Carney, MS, Caren Kamberg, MSPH, Ellen Reinisch, MS, and Marjorie J. Sherwood, MD, were key contributors to project development and implementation. Andrea Steiner, MS, gave valuable editorial advice. Our policy advisory board (John C. Beck, MD, Barbara J. Burns, PhD, Monroe T. Gilmour, MD, Paul F. Griner, MD, Charlene Harrington, RN, PhD, T. Reginald Harris, MD, Rosalie Kane, DSW, Shirley Kellie, MD, Judith R. Lave, PhD, Charles E. Lewis, MD, Joseph Martin, Francis D. Moore, Jr, MD, Richard N. Pierson, Jr, MD, and James F. Rodgers, PhD) provided astute advise and guidance. We acknowledge the important contributions of the 297 hospitals whose medical records we reviewed. Without the efforts of these individuals and institutions, this evaluation could not have been successfully completed. Finally, we recognize and thank Florence McGinty, without whose secretarial skills this article would not have been produced.

References

1. Prospective Payment Assessment Commission. *Medicare Prospective Payment and the American Health Care System: Report to Congress.* Washington, DC; February 1986.
2. Department of Health and Human Services. *Hospital Prospective Payment for Medicare: Report to Congress.* Baltimore, Md; December 1982.
3. Guterman S, Eggers PW, Riley G, et al. The first three years of Medicare prospective payment: an overview. *Health Care Financing Rev.* 1988; 9:67-77.
4. Draper D, Kahn KL, Reinisch EJ. Studying the effects of the DRG-based prospective payment system on quality of care: design, sampling, and fieldwork. *JAMA.* 1990;264:1956-1961.
5. 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.
6. 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.
7. Rubenstein LV, Kahn KL, Reinisch EJ, et al. Changes in quality of care for five diseases as measured by implicit review, 1981-1986. *JAMA.* 1990; 264:1974-1979.
8. Kosecoff J, Kahn KL, Rogers WH, et al. Prospective payment system and impairment at discharge: the 'quicker-and-sicker' story revisited. *JAMA.* 1990;264:1980-1983.
9. 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.
10. Rogers WH, Draper D, Kahn KL, et al. Quality of care before and after implementation of the DRG-based prospective system: a summary of effects. *JAMA.* 1990;264:1989-1994.
11. DesHarnais S, Chesney J, Fleming S. Trends and regional variations in hospital utilization and quality during the first two years of the prospective payment system. *Inquiry.* 1988;25:374-382.
12. Mayer-Oakes SA, Oye RK, Leake B, Brook RH. The early effect of Medicare's prospective payment system on the use of medical intensive care services in three community hospitals. *JAMA.* 1988;260:3146-3149.
13. Sager MA, Leventhal EA, Easterling DV. The impact of Medicare's prospective payment system on Wisconsin nursing homes. *JAMA.* 1987;257: 1762-1766.
14. Sager MA, Easterling DV, Kindig DA, Anderson OW. Changes in the location of death after passage of Medicare's prospective payment system. *N Engl J Med.* 1989;320:433-439.
15. Fitzgerald JF, Fagan LF, Tierney WM, et al. Changing patterns of hip fracture care before and after implementation of the prospective payment system. *JAMA.* 1987;258:218-221.
16. Fitzgerald JF, Moore PS, Dittus RS. The care of elderly patients with hip fracture. *N Engl J Med.* 1988;319:1392-1397.
17. Palmer RM, Saywell RM Jr, Zollinger TW, et al. The impact of the prospective payment system on the treatment of hip fractures in the elderly. *Arch Intern Med.* 1989;149:2237-2241.
18. Gerety MB, Soderholm-Difatte V, Winograd CH. Impact of prospective payment and discharge location at the outcome of hip fracture. *J Gen Intern Med.* 1989;4:388-391.
19. Kahn KL, Brook RH, Draper D, et al. Interpreting hospital mortality data: how can we proceed? *JAMA.* 1988;260:3625-3628.
20. Gruppo Italiano per lo Studio Della Streptochi-Nachi Nell Infarcto Miocardico. Long-term effects of intravenous thrombolysis in acute myocardial infarction: final report of the GISSI study. *Lancet.* 1987;2:871-874.
21. *1986 Medicare Provider Analysis Record (MedPAR) File.* Baltimore, Md: Health Care Financing Administration; 1986.
22. Vladeck BC. Hospital prospective payment and the quality of care. *N Engl J Med.* 1988; 319:1411-1413.
23. Iglehart JK. Early experience with prospective payment of hospitals. *N Engl J Med.* 1986; 314:1460-1464.
24. Prospective Payment Assessment Commission. *1989 Adjustments to the Medicare Prospective Payment System: Report to Congress.* November 1988:B-3.
25. Witsberger C, Kominsky G. *Recent Trends in Length of Stay for Medicare Surgical Patients.* Santa Monica, Calif: RAND Corp; 1990.

Studying the Effects of the DRG-Based Prospective Payment System on Quality of Care

Design, Sampling, and Fieldwork

David Draper, PhD; Katherine L. Kahn, MD; Ellen J. Reinisch, MS; Marjorie J. Sherwood, MD; Maureen F. Carney, MS; Jacqueline Kosecoff, PhD; Emmett B. Keeler, PhD; William H. Rogers, PhD; Harry Savitt, PhD; Harris Allen, PhD; Kenneth B. Wells, MD, MPH; David Reboussin, MS; Robert H. Brook, MD, ScD

We have conducted a nationally representative before-after study of the effects of the diagnosis related groups–based prospective payment system (PPS) on quality of in-hospital care for aged Medicare patients. We used a pre-post design with multiple time points in both the pre-PPS (calendar years 1981 and 1982) and post-PPS (July 1985 through June 1986) periods. We gathered clinically detailed data from medical records of patients with one of six diseases and supplemented these data with postdischarge information from Health Care Financing Administration files. We used a stratified multistage cluster sampling design with data gathered on 16 758 patients chosen from 297 hospitals in 30 areas in five states. Our hospital participation rate was 97%; we successfully accessed 96% of the medical records we requested; and our mean item-level reliability score was 0.80. Our sample matches the nation closely on hospital urbanicity, size, teaching status, ownership, and percentages of Medicare and Medicaid patients, and patient demographics and mortality.

(*JAMA*. 1990;264:1956-1961)

From the Health Program of the RAND Corp, Santa Monica, Calif (Drs Draper, Kahn, Sherwood, Keeler, Rogers, Allen, Wells, and Brook, Mss Reinisch and Carney, and Mr Reboussin); the RAND Graduate School of Policy Studies, Santa Monica, Calif (Drs Draper, Keeler, Rogers, and Brook); the Departments of Medicine (Drs Kahn, Kosecoff, and Brook), Health Services (Drs Kosecoff and Brook), and Psychiatry and Biobehavioral Sciences (Dr Wells), UCLA; Value Health Sciences Inc, Santa Monica, Calif (Dr Kosecoff); and the Office of Research and Demonstrations of the Health Care Financing Administration, Baltimore, Md (Dr Savitt).

The opinions, conclusions, and proposals in this article are those of the authors alone and do not necessarily represent the views of any of these organizations, the RAND Corp, or UCLA.

Reprint requests to the RAND Corp, 1700 Main St, PO Box 2138, Santa Monica, CA 90406-2138 (Dr Draper).

IN 1983, THE Health Care Financing Administration (HCFA) changed the way hospitals were reimbursed for treating patients under the Medicare program.[1] Prior to 1983, hospitals received payment for all services provided, subject to appropriateness review. Since 1983, under the prospective payment system (PPS), hospitals have been paid an amount based largely on flat rates per admission calculated for each of approximately 470 diagnosis related groups. Because the new payment system contains incentive to decrease length of stay and substitute lower-cost services and procedures, concern has arisen that the quality of health care may have declined.

Since 1985, we have been conducting a national study to investigate the effects of the PPS on quality of care for hospitalized Medicare patients. Other articles in this series present our findings.[2-7] In this article, we summarize our design and sampling decisions, give details on the fieldwork involved in gathering our primary data, and present results on the composition and national representativeness of our final sample.

METHODS

Choice of Treatment, Control Groups, and Study Years

The PPS was not introduced in 1983 as a controlled experiment. Instead, before October 1983, hospitals were reimbursed for treating Medicare patients under the old retrospective payment system, and during the year from October 1983 to September 1984 nearly all acute care general hospitals were phased into prospective reimbursement. The exceptions were hospitals in the waiver states—Maryland, Massachusetts, New Jersey, and New York—where reimbursement alternatives to the PPS were used until 1986.

Because the PPS was introduced in

Reprinted from *JAMA, The Journal of the American Medical Association*, October 17, 1990, Vol. 264, No. 15, pp. 1956–1961, © 1990, American Medical Association. Reprinted by permission.

this nonexperimental way, a prospective controlled trial evaluating its effects on quality of care was not possible. We instead designed a retrospective observational study, in which we contrasted data on Medicare patients prior to and subsequent to 1983. We considered supplementing such data with pre- and post-PPS information from a control group (eg, patients from the waiver states, non–Medicare patients aged 55 to 64 years in the PPS states, or patients from another country such as Canada), but funding limitations prevented this.

Because pre-post designs may be confounded by secular trends over time, we gathered data at multiple time points in both the pre- and post-PPS periods, so that such trends might be at least roughly estimated. We report only differences pre- and post-PPS in the middle five articles in this series[2-6] and discuss issues of trend estimation and causality in the final article in the series.[7]

We chose as our pre-PPS sampling window the period from January 1, 1981, to December 31, 1982, and as our post-PPS window the period from July 1, 1985, to June 30, 1986. Our design concentrated 50% of the sampling in each of the pre- and post-PPS periods: 20% in 1981, 30% in 1982, and 25% in each of 1985 and 1986.

Our design was longitudinal at the hospital level and cross-sectional at the patient level; ie, we gathered data at each of our sampled hospitals in all study years, with different patient cohorts sampled in each time period within the chosen hospitals. The longitudinal nature of the hospital sampling increased the accuracy of the pre-post comparison by holding the hospital factor constant.

We have elsewhere[8] advocated the measurement of quality of care in a disease-specific way. However, instrumentation costs limit the number of diseases that may be studied in detail. After consultation with an expert panel, we selected six diseases for study: congestive heart failure, acute myocardial infarction, pneumonia, cerebrovascular accident, hip fracture, and depression. We discuss the extent to which our findings from these six diseases generalize to other Medicare patients in another article in this series.[9]

Summary of Sampling Plan

We used a stratified, multistage cluster sampling plan with four levels of sampling hierarchy: states, areas within states, hospitals within areas, and patients within hospitals. We oversampled hospitals treating many Medicare patients and chose approximately the same number of patients from each hospital in a way that produced patient-level national representativeness. In our sampling design we chose five states, with four to eight areas per state for a total of 30 areas nationwide; six to 18 hospitals per area for a total of approximately 60 hospitals per state and 300 hospitals overall; and about 57 patients per hospital for a total of approximately 17 000 patients.

Choice of States.—We selected our final states purposively, subject to eligibility criteria and stratification goals. Eligibility considerations excluded the waiver states and states with either too few hospitals or a mixture of hospitals that was either too urban or too rural. The main stratification goal in the choice of states was geographic diversity, with one state from each region of the country. Our final sample included California, Texas, Indiana, Pennsylvania, and Florida.

Choice of Areas. — Photocopying sampled medical records and sending the copies to a central location in the chosen states for abstraction would have been desirable on cost grounds, but the photocopy quality of microfilmed records, which made up a nontrivial portion of the pre-PPS sample, was too poor to permit this option. Given the resulting cost restrictions on data collector travel, the only feasible sampling plan involved dividing the sampled states into geographic areas and clustering the chosen hospitals in a sample of these areas.

We used geographic diversity within state and six hospital-level factors as stratification variables in our final area choice: urbanicity, percentages of Medicare and Medicaid patients, size, teaching status, and hospital ownership (eg, proprietary vs nonprofit). We considered a large number of designs, each with a total of 20 to 30 areas, by conducting a computer-aided search among all possible choices of four to eight areas in each of our five states. Our final choice was purposive and had 30 areas, with four to eight areas per state.

Choice of Hospitals.—We based the hospital-level sampling frame on the 1984 HCFA Provider of Services file.[10] We used three eligibility criteria at the hospital level. First, we restricted sampling to short-term acute care facilities and excluded veterans', military, and psychiatric hospitals. Second, we also restricted attention to hospitals that were in existence during the entire period from 1981 to 1986. Third, a small number of hospitals in our chosen states had 15 or fewer patients per year with one or more of our six study diseases (based on HCFA's 1984 MedPAR file[10]) and were judged to be too small.

The final hospitals were chosen by defining a stratification grid indexed by size, urbanicity, and hospital poverty status. We defined "high-poverty hospitals" as those facilities whose percentage of Medicaid patients was at or above the 90th percentile of the Medicaid distribution in the state in which the hospital was located. Our goal was representativeness with respect to size and urbanicity and an oversampling by a factor of 2 of both high-poverty hospitals and city-county facilities. Hospitals were then chosen by a restricted randomization procedure that maximized representativeness with respect to ownership, teaching intensity, and percentage of Medicare patients, while varying the number of hospitals per area from 6 to 18 in such a way that our stratification targets were achieved. In this manner, 300 hospitals were chosen.

Choice of Patients.—We based the patient-level sampling frame on lists, generated by the chosen hospitals, of all patients hospitalized in each study year with one of the study diseases, as indicated by *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) codes (Table 1). Twelve hospitals were unable to provide such lists; we built the patient sampling frames for those hospitals from HCFA's MedPAR file. We chose simple random samples of patients in the cells of a three-by-six stratification grid indexed by study period and disease. The goal was balance across all diseases, with nine or 10 patients chosen from each disease category in each hospital, distributed within each disease across study periods in approximate proportions of 20% (1981), 30% (1982), and 50% (1985 and 1986).

To meet our goal of 17 000 total patients, we examined a somewhat larger number of records because some records were unavailable for study and others did not meet our clinical inclusion criteria (see below). In total, 22 795 records were requested from the participating hospitals.

Data Collector Training, Abstraction, and Monitoring

Nurses and medical records personnel who were experienced with clinical data were selected by the professional review organizations in the chosen states to become data collectors. After demonstrating adequate skills, 52 data collectors participated in training that lasted 17 days, with at least 2 days focused exclusively on the study of each of our six diseases.

During the period of data collection,

Table 1.—*ICD-9-CM** Codes Defining the Six Study Diseases

Disease	*IDC-9-CM* Codes†
Congestive heart failure	428.*xx*
Acute myocardial infarction	410.*xx*
Pneumonia	480.9*x*
	481.*xx*
	482.*xx*
	483.*xx*
	485.*xx*
	486.*xx*
	487.0*x*
	507.0*x*
	510.*xx*
	511.1*x*
	513.0*x*
	799.1*x*
Cerebrovascular accident	431.*xx*
	434.*xx*
	436.*xx*
Hip fracture	820.0*x*‡
	820.2*x*
	820.8*x*
Depression	296.2*x*
	296.3*x*
	298.0*x*
	300.4*x*
	309.0*x*
	309.1*x*
	311.*xx*

**ICD-9-CM* indicates *International Classification of Diseases, Ninth Revision, Clinical Modification.*
†*x* = any value 0 to 9 or no value.
‡Does not include 820.01.

we monitored abstractors with a series of interrater reliability studies (see below) and gave feedback to data collectors who were identified as having problems. A project manager who was familiar with the hospital in which the abstraction took place supervised the work. In addition, each completed abstraction form was reviewed by both a physician and a nurse for an average of 30 minutes per record to assess internal consistency and to assure that coding was consistent with supporting clinical data. Certain types of clinical information from the medical record (eg, the exact words characterizing a patient's stroke if one occurred) were written verbatim into the abstraction form, and specified pieces of data (eg, chest roentgenogram reports, admission histories and physicals, and discharge summaries) were photocopied and attached to the form. Discrepancies that could not be resolved during the review process were returned for reabstraction. Physicians interpreted photocopies of roentgenogram reports and some electrocardiogram tracings and reports.

To maintain the confidentiality of hospitals, patients, and physicians, coded identifiers were assigned to hospitals and patients. No identifying information about physicians was obtained.

Each disease - specific abstraction form[11-16] contained approximately 700 items grouped into five categories: inclusion criteria, demographics and sickness at admission, explicit information about the processes of care given to the patient, patient outcomes, and patient status at discharge. Examples include physician documentation of preadmission symptoms of myocardial infarction, frequency of nurses' blood pressure readings on hospital day 2 (congestive heart failure), prehospital mental status (pneumonia), use of coumarin at any time during hospitalization (cerebrovascular accident), and number of physical therapy sessions on postoperative day 1 (hip fracture). More details about data elements can be found in the other articles in this series.[2-7] The abstraction forms took an average of 90 minutes per record to complete.

Interrater Reliabilities

We assessed interrater reliabilities for each disease using both records that the data collectors knew to be test cases and records that they did not know were being monitored. We used κ scores[17] to measure by how much the agreement between different readings of the same medical record exceeded chance.

Inclusion Criteria

We used inclusion criteria to assure a homogeneous group of patients with the chosen diseases in the pre- and post-PPS periods. To be included in one of our disease-specific samples, patients had to be at least 65 years of age, admitted during one of our study years, and hospitalized for the indicated disease. Patients with the study disease as a complication of hospitalization, rather than as a reason for admission, were not eligible for the sample.

Secondary Data Collection

In addition to the clinical primary data previously described, we also collected secondary data of two types: (1) information about postdischarge outcomes that was merged with the records in our primary database in order to examine the effects of the PPS on outcomes during the year following hospitalization, and (2) national data on mortality trends in our study diseases from 1980 to 1986 with which to compare our sample data for validation purposes.

We compiled three kinds of postdischarge outcome information: mortality, hospital readmissions, and nursing home stays. To obtain postdischarge mortality data, we used patients' last names, first names, dates of birth, and health insurance claim numbers from the medical record to match our sampled patients with the corresponding records in HCFA's health insurance master file.[10]

To validate our sampling with respect to mortality, we obtained data from HCFA's MedPAR file on all patients hospitalized with our five nonpsychiatric diseases (n = 2 062 610) in the 27-quarter period from January 1980 to September 1986. Variables we extracted from the MedPAR file included death status within 30 days of admission, gender, age, and diagnosis related group. We established a correspondence between our *ICD-9-CM*–based disease definitions (Table 1) and diagnosis related groups and used these data to calculate age- and sex-specific and age- and sex-adjusted 30-day death rates for each of our study diseases, and in the aggregate across our study diseases, in each of the 27 quarters.

We then used the postdischarge mortality data on our sampled patients to compute age- and sex-adjusted quarterly sample mortality rates in the 30-day period following admission, by disease and in the aggregate across diseases. To compare these data with the national MedPAR mortality data previously described, we plotted the national age- and sex-adjusted 30-day quarterly mortality series on the same graph with the quarterly 30-day mortality data from our study sample and marked off error bars of 2 SEs either way from the observed sample mortality series. We estimate that noncomparabilities between national and sample mortality values arising from our use of inclusion criteria were small.

Effects of Sampling Plan on Analysis

The sampling plan we employed had four features requiring special attention during the analysis:

1. Our oversampling of patients from hospitals serving an unusually large fraction of Medicaid patients would yield somewhat biased raw findings if the quality of care in these facilities differed substantially from that in other hospitals. To arrive at nationally representative findings, it was necessary to reweight our raw patient-level results, giving less weight to facilities serving an unusually large fraction of Medicaid patients.[18] The weighted and unweighted results differed little in most of our major analyses. In what follows in this and other articles in this series,[2-7] we present unweighted findings unless otherwise indicated.

2. The clustering of our sampled patients in only five states, 30 areas, and 297 hospitals has implications for the accuracy of our results. If there is more similarity on average between two patients in the same state, area, or hospital in the care they receive than between two patients in different states, areas, or hospitals, and no adjustment for this intracluster correlation[19] is made, the result will be an overstatement of the precision of our findings. To

adjust for clustering and to produce SEs that accurately reflect the information content of our sample, we (1) calculated provisional SEs for all of our major estimates, as if we had gathered our data using simple random sampling, (2) computed "inflation factors" based on intracluster correlations that measured the amount of information in our sample relative to that obtained in a simple random sample of the same size, and (3) multiplied our provisional SEs by the inflation factors, thereby adjusting the significance of our results downward. The inflation factors ranged from 1.1 for outcomes such as in-hospital and 30-day death to 2.9 for our composite score, aggregating across diseases, that summarizes the quality of the processes of in-hospital care. The latter inflation factor, which is based on an unusually high intracluster correlation of .16, indicates a remarkable degree of homogeneity within hospitals in the processes of care rendered.

3. The longitudinal nature of the hospital sampling, in which we gathered data from all of our study hospitals in both the pre- and post-PPS periods, also had implications for the accuracy of our results. Holding the hospital factor constant acted on the precision of our findings in a manner opposite to that of the clustering: the latter decreased accuracy, when compared with independent simple random sampling in each of the pre- and post-PPS periods, while the former increased accuracy because of positive correlations over time in the patterns of care given to patients within hospitals. We computed "deflation factors" to adjust simple random sampling SEs for the longitudinal hospital sampling and found that the clustering and longitudinal effects approximately canceled each other in our analyses. Thus, we were able to analyze the data essentially as if they had been gathered with simple random sampling in a cross-sectional fashion.

4. We present results in this series of articles[2-7] both at the disease-specific level and in the aggregate across diseases. Two issues arise in producing across-disease summaries: reweighting the diseases back to their actual frequencies in the Medicare population, and properly accounting for the degree to which pre-post differences themselves differ by disease (ie, accounting for interactions between the pre-post and disease factors). Reweighting by disease prevalence is potentially necessary because we took samples of roughly equal size in each of our six study diseases, even though congestive heart failure is more than three times more frequent among Medicare patients than hip fracture, for example.

We computed weights necessary to reweight our raw findings back to the population of all Medicare patients with one of our six study diseases and did sensitivity analyses to see how much this reweighting affected our results. In all of the cases we examined involving our major study findings, the weighted and unweighted results differed little. We therefore report unweighted findings unless otherwise noted. Regarding pre-post by disease interactions, we tried not to emphasize across-disease summaries when the interactions were large (ie, when there was serious disagreement among our diseases in the size of the difference pre- and post-PPS), and we tried to conservatively report the significance of the pre-post PPS effect, when the interactions were small to moderate, by adjusting the pre-vs post-PPS significance in a manner consistent with the size of the interaction.[20]

Table 2.—Percentages of Records Excluded by Disease and Time Period*

	1981	1982	1985 and 1986	Total
Disease				
Congestive heart failure	29.8	28.4	19.7	24.5
Acute myocardial infarction	18.7	16.7	14.8	16.2
Pneumonia	39.3	38.1	35.9	37.3
Cerebrovascular accident	22.4	21.8	13.1	17.7
Hip fracture	21.8	18.3	14.9	17.4
Depression	21.7	25.9	24.2	24.4
Total No. (%) excluded	1266 (26.2)	1677 (25.5)	2244 (21.2)	5187 (23.6)
Total No. included	3455	4949	8354	16 758
Total No. (Excluded + Included)	**4721**	**6626**	**10 598**	**21 925**

*The percentages given are of the following form: number of excluded records/(number of included + excluded records).

Table 3.—Frequent Reasons for Record Exclusion by Disease

	Diseases*						
	CHF	AMI	PNE	CVA	HIP	DEP	6 Diseases
Frequent reasons							
Not hospitalized during study year, %	0.3	0.4	0.2	0.3	0.3	0.2	0.3
No eligible *ICD-9-CM*† code, %	1.0	1.5	1.6	1.2	1.5	2.5	1.6
Age <65 y, %	0.2	0.3	0.5	0.3	0.1	0.7	0.4
Transferred from another acute care hospital, %	1.8	4.1	1.8	3.3	2.2	ND‡	2.6
No symptoms/signs of disease at admission, %	16.8	6.5	27.5	11.0	2.2	10.9	12.5
AMI at admission, %	0.7	ND	0.2	0.3	0.2	ND	0.4
Other competing conditions,§ %	1.3	2.2	3.6	0.6	10.3	6.4	4.1
Poor prognosis or active cancer, %	1.8	ND	ND	ND	ND	1.2	1.5
Total No. excluded	916	550	1632	608	581	880	5167
Total No. included	2824	2853	2749	2824	2762	2746	16 758
Total No. (Excluded + Included)	**3740**	**3403**	**4381**	**3432**	**3343**	**3626**	**21 925**

*CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; HIP, hip fracture; and DEP, depression.

†*ICD-9-CM* indicates *International Classification of Diseases, Ninth Revision, Clinical Modification.*

‡ND indicates that the disease-specific abstraction form did not specify this as an exclusion criterion; as a result, no data were collected on this item.

§Other competing conditions included important clinical problems present at admission that were likely to influence process and/or outcome in ways not seen with patients without these conditions (eg, the presence of bilateral hip fracture or multiple trauma were exclusion criteria for hip fracture patients).

RESULTS

Sampling

Almost all of the 300 selected hospitals agreed to participate in the study. Five hospitals refused and were replaced by hospitals in the same area with similar hospital and patient characteristics. Three other hospitals also refused to participate and were not replaced, resulting in a final sample of 297 hospitals and a hospital-level participation rate of 97% (297/305).

Our final sample had 51 to 62 patients per hospital (except for 12 small hospitals, where patients per hospital ranged from 17 to 50), with an average of approximately 57 patients per hospital and a total of 16 758 patients, slightly fewer than our target of 17 000.

Hospitals were able to find the medi-

cal records we requested in almost all cases. Of the 22 795 medical records requested, 870 (3.8% of the requested records) were not available for review, leaving 21 925 total records reviewed. The number that were not available was higher in 1981 (6.6% of those requested) and 1982 (5.0%) than in 1985 and 1986 (1.7%). The numbers of records unavailable were sufficiently small that any bias arising from their unavailability would be too small to significantly change the study's major findings.

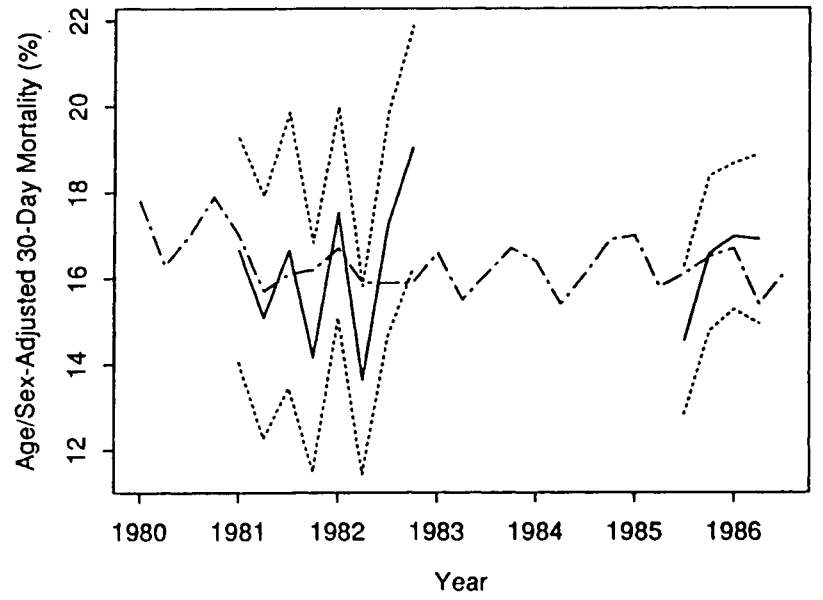
Of the 21 925 records reviewed, 5167 (24%) were excluded (Table 2). The fraction of records excluded in the post-PPS period (21%) was significantly lower than the fraction excluded in the pre-PPS period (26%; $P<.01$). Table 3 gives the most frequent causes of exclusion. The most common reason was that clinically detailed review of the medical record showed that the patient did not have symptoms or signs of the disease suggested by the hospital's assignment of *ICD-9-CM* codes. For example, 17% of the patients with a congestive heart failure code on their medical records did not have signs of congestive heart failure at admission, ie, they did not have either chest roentgenogram evidence for heart failure or leg edema at admission.

We were able to obtain accurately merged postdischarge mortality information from the HCFA's files on 92% of our sample (12 821 of 14 012 patients).

Table 4.—Detectable Pre-Post Differences: Process* and Outcome

Disease	Outcome: 30-Day Mortality, Adjusted for Sickness at Admission		
	Pre-PPS† Rate,%	Pre-Post Change Detectable With 80% Power, Percentage Points	SE for Estimated Pre-Post Difference, Percentage Points
Congestive heart failure	14.7	3.4	1.3
Acute myocardial infarction	24.4	4.2	1.5
Pneumonia	15.9	3.6	1.3
Cerebrovascular accident	21.3	4.0	1.4
Hip fracture	5.1	2.1	0.8
Aggregating across diseases	16.7	1.6	0.6

*Process: the study's power to find 1.1-point or greater disease-specific pre-post differences on a 100-point process scale with a patient-level SD of 10 was at least 80%. In the aggregate across diseases, 0.5-point or larger differences on such a scale were detectable with at least 80% power.
†PPS indicates prospective payment system.



Quarterly 30-day mortality trend for five diseases. Solid line indicates sample values (n = 12 823); dashed lines, national values (n = 2 062 610); and dotted lines, 2 SEs either way from sample values.

Interrater Reliabilities

For 10 reliability records known by the data collectors to be test cases, each rated by an average of 47 different data collectors, we found across-disease item-level κ scores averaging 0.86. For 162 different records rated by two different data collectors who did not know they were being monitored, we obtained across-disease item-level κ scores averaging 0.78. Overall, the 10th percentile of the distribution of our κ scores was 0.63, the 90th percentile of our κ scores was 1.0; 5.6% of our κ scores were below 0.4 (a level generally recognized as signifying poor reliability[17]), and 83% were above 0.75 (a level implying excellent reliability[17]).

Effects Detectable With the Study's Sampling Resources

Table 4 gives examples of the differences from pre- to post-PPS that were

Table 5.—Characteristics of the Study Sample by Disease and Time Period*

Patient Characteristic	Study Disease†												
	CHF		AMI		PNE		CVA		HIP		Total		
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Total
≥80 y, %	41	41	25	27	41	42	41	41	58	58	41	42	41
Male, %	43	45	53	50	51	51	45	44	21	23	43	43	43
Nonwhite, %	23	22	18	16	19	20	22	20	14	13	19	18	19
Medicaid, %	16	15	12	12	19	19	14	13	17	15	15	15	15
Preadmission residence at nursing home, %	8	9	4	5	22‡	26‡	12	11	24‡	20‡	14	14	14
Sample size	1359	1465	1416	1437	1341	1408	1382	1442	1358	1404	6856	7156	14 012

*The results were reweighted to achieve national representativeness.
†CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, penumonia; CVA, cerebrovascular accident; and HIP, hip fracture.
‡Difference pre- vs post-PPS is significant ($P<.01$).

detectable with the study's resources. It can be seen from this table that small disease-specific and aggregate differences in the average quality of processes of care, and a modest difference in aggregate mortality across all five diseases, stood a high chance of being found, while disease-specific differences in mortality would have had to have been fairly large to be detectable with high likelihood. The estimates in the table are typical of the power conclusions to be drawn about the many quantitative and qualitative response variables in our study: our data resources were sufficient to find small disease-specific differences for continuous outcomes and moderate differences aggregating across diseases for dichotomous outcomes.

Effects smaller than those referred to in Table 4 as detectable with 80% power may actually be found to be significant by the study, as can be seen from the column giving SEs for the estimated differences pre- and post-PPS in 30-day mortality, adjusted for sickness at admission.

Representatives of Final Sample

After reweighting, our sample matched the nation closely with respect to hospital size, urbanicity, percent of Medicare and Medicaid patients, teaching intensity, and ownership. With each of these variables divided into three to six categories, the largest discrepancy between national and reweighted sample prevalences in any category was less than 2 percentage points, and national and reweighted sample means and SDs for these variables (when relevant) agreed to within less than 1%.[18]

The Figure compares our sampled age- and sex-adjusted 30-day mortality values with known national values on a quarterly basis, in the aggregate across diseases; the results were similar at the disease-specific level. The national values fell within the sample 95% confidence limits in 11 of the 12 quarters, and the pattern of positive and negative deviations of the sampled values from the national mortality figures revealed no bias. Both of these observations are consistent with the hypothesis that our sample is representative of the nation with respect to 30-day mortality.

Composition of Final Patient Sample

Table 5 gives various characteristics of the study's patient sample, both before and after the PPS, by disease and in the aggregate across diseases. Demographics remained stable from pre- to post-PPS, but there was some change in the fraction of patients admitted from a nursing home; this percentage rose from 22% to 26% pre- to post-PPS for pneumonia patients but fell by the same amount (from 24% to 20%) for hip fracture patients ($P<.01$).

SUMMARY

In this article we have presented the design and sampling choices that led to our collection of data on 8404 Medicare patients treated before the introduction of the PPS and 8354 patients treated after the PPS was implemented, and we have demonstrated the national representativeness of these patient samples with respect to a number of key hospital- and patient-level variables. Other articles in this series[2-7] present the results of our analysis of these data.

This study was supported by cooperative agreement 18-C-98853/9-03 issued by the Health Care Financing Administration of the US Department of Health and Human Services, Baltimore, Md. The entire RAND costs ($3.9 million) of this project were funded with federal money.

We gratefully acknowledge the collaborative effort of the five professional review organizations that enabled this work to be completed. Participants from these organizations included medical directors, physician consultants and reviewers, project directors, and review coordinators from each of the five study states. In particular, we appreciate the keen clinical insight of the professional review organizations' physician specialists with whom we consulted throughout this study. Stanley S. Bentow, MS, and Caren Kamberg, MSPH, were key contributors to project development and implementation. Andrea Steiner, MS, gave valuable editorial advice. We thank Paul Eggers, PhD, who helped us acquire data for validation purposes that would otherwise have been impossible to obtain. Our policy advisory board (John C. Beck, MD, Barbara J. Burns, PhD, Monroe T. Gilmour, MD, Paul F. Griner, MD, Charlene Harrington, RN, PhD, T. Reginald Harris, MD, Rosalie Kane, DSW, Shirley Kellie, MD, Judith R. Lave, PhD, Charles E. Lewis, MD, Joseph Martin, Francis D. Moore, Jr, MD, Richard N. Pierson, Jr, MD, and James F. Rodgers, PhD) provided astute advice and guidance. We acknowledge the important contributions of the 297 hospitals whose medical records we reviewed. Without the efforts of these individuals and institutions, this evaluation could not have been successfully completed.

References

1. Prospective Payment Assessment Commission. *Medicare Prospective Payment and the American Health Care System: Report to Congress*. Washington, DC: US Government Printing Office; 1986.
2. 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.
3. 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.
4. Rubenstein LV, Kahn KL, Reinisch EJ, et al. Changes in quality of care for five diseases as measured by implicit review, 1981 to 1986. *JAMA*. 1990;264:1974-1979.
5. 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.
6. Kosecoff J, Kahn KL, Rogers WH, et al. Prospective payment system and impairment at discharge: the 'quicker-and-sicker' story revisited. *JAMA*. 1990;264:1980-1983.
7. Rogers WH, Draper D, Kahn KL, et al. Quality of care before and after implementation of the DRG-based prospective system: a summary of effects. *JAMA*. 1990;264:1989-1994.
8. Kahn KL, Brook RH, Draper D, et al. Interpreting hospital mortality data: how can we proceed? *JAMA*. 1988;260:3625-3628.
9. 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;264:1969-1973.
10. Bureau of Data Management and Strategy. *HCFA Data Profiles*. Baltimore, Md: Health Care Financing Administration; 1986.
11. 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 Corp; 1988.
12. 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 Corp; 1988.
13. Roth CP, Kahn KL, Sherwood MJ, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Pneumonia*. Santa Monica, Calif: RAND Corp; 1988.
14. 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 Corp; 1988.
15. Sherwood MJ, Kahn KL, Kosecoff J, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Hip Fractures*. Santa Monica, Calif: RAND Corp; 1988.
16. Wells KB, Kosecoff J, Sherwood MJ, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Depression*. Santa Monica, Calif: RAND Corp; 1988.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
18. Kahn KL, Draper D, Keeler EB, et al. *The Effects of the DRG-Based Prospective Payment System on Quality of Care for Hospitalized Medicare Patients: Supplemental Design, Methods, and Results*. Santa Monica, Calif: RAND Corp; 1990.
19. Cochran WG. *Sampling Techniques*. 3rd ed. New York, NY: John Wiley & Sons Inc; 1977.
20. Scheffé H. *The Analysis of Variance*. New York, NY: John Wiley & Sons Inc; 1959.

Changes in Sickness at Admission Following the Introduction of the Prospective Payment System

Emmett B. Keeler, PhD; Katherine L. Kahn, MD; David Draper, PhD; Marjorie J. Sherwood, MD; Lisa V. Rubenstein, MD, MSPH; Ellen J. Reinisch, MS; Jacqueline Kosecoff, PhD; Robert H. Brook, MD, ScD

We developed disease-specific measures of sickness at admission based on medical record data to study mortality of Medicare patients with one of five conditions (congestive heart failure, acute myocardial infarction, cerebrovascular accident, pneumonia, and hip fracture). We collected an average of 73 sickness variables per disease, but our final sickness-at-admission scales use, on average, 19 variables. These scales are publicly available, and explain 25% of the variance in 30-day postadmission mortality for patients with acute myocardial infarction, pneumonia, or cerebrovascular accident. Sickness at admission increased following the introduction of the prospective payment system (PPS). For our five diseases combined, the 30-day mortality to be expected because of sickness at admission was 1.0% higher in the 1985-1986 period than in the 1981-1982 period (16.4% vs 15.4%), and the expected 180-day mortality was 1.6% higher (30.1% vs 28.5%). Studies of the effects of PPS on mortality must take this increase in sickness at admission into account.

(*JAMA*. 1990;264:1962-1968)

THE INTRODUCTION in 1983 of the prospective payment system (PPS) has raised clinical and policy questions. Did the change in financial incentives from a cost-plus reimbursement system to a fixed-price system result in fewer sick patients' being admitted to the hospital? Did mortality within 30 or 180 days following hospitalization change, after adjusting for sickness at admission? Measuring how sick a patient is at the time of hospital admission is a prerequisite for answering such questions. In this report we present measures of sickness at admission for five diseases: congestive heart failure, acute myocardial infarction, pneumonia, cerebrovascular accident, and hip fracture. These measures can also be used to aid clinical decisions at the individual patient level and help hospitals monitor their outcomes. After describing the measures, we present data that show how sickness at admission changed after the introduction of prospective payment.

From the Health Program of the RAND Corp, Santa Monica, Calif (Drs Keeler, Kahn, Draper, Sherwood, Brook, and Rubenstein and Ms Reinisch); the Department of Medicine (Drs Kahn, Brook, Rubenstein, and Kosecoff) and Department of Health Services (Drs Brook and Kosecoff), UCLA School of Medicine; and Value Health Sciences Inc, Santa Monica, Calif (Dr Kosecoff).

The opinions, conclusions, and proposals in the text are those of the authors alone and do not necessarily represent the views of any of these organizations, the RAND Corp, or UCLA.

Reprint requests to the RAND Corp, 1700 Main St, PO Box 2138, Santa Monica, CA 90406-2138 (Dr Keeler).

PATIENTS AND METHODS

The sample for this study includes 14 012 patients who are aged 65 years or older. We present details of the sampling design and inclusion criteria elsewhere.[1,2]

Variables That Measure Sickness at Admission

In developing our measures of sickness at admission, we used previously published severity measures, including those developed to predict death for patients admitted to intensive care units[3-5] and measures developed for specific diseases.[6-9] We used literature review, clinical judgment, and disease-specific consensus panels to identify other variables that have been considered important clinical predictors of the outcomes we chose to study.[2]

We used disease-specific abstraction forms to collect data about sickness at admission from the medical records of hospitalized patients.[10-14] We collected data about acute and chronic morbid and comorbid diseases, function, the number of body systems with pathologic findings, and the APACHE II (*A*cute *P*hysiology *a*nd *C*hronic *H*ealth *E*valuation) Acute Physiological Score (APS) variables.[3] For patients who had an acute myocardial infarction, we also collected data used to score the Killip Scale and the Norris Coronary Prognostic Index.[6,7] For patients who had hip fractures, we collected data for the Goldman Preoperative Risk Index.[8] We always collected the first available data. If the patient had missing data on day 1, we accepted data from day 2, because day 2 values may represent the admission status of patients who were admitted late at night.

Reprinted from *JAMA, The Journal of the American Medical Association*, October 17, 1990, Vol. 264, No. 15, pp. 1962-1968, © 1990, American Medical Association. Reprinted by permission.

Details on the full set of variables collected, analyses underlying the scales, and examples of how to use the scales to predict mortality are available.[15]

Analytic Methods

Modeling decisions (scaling of important variables and subsequent variable selection) were based on a two-thirds random sample of the patients with each disease. We validated our modeling choices by repeating the modeling effort on one third of the patients and on five bootstrap[16] replicates of the data sets. Final weights were estimated on the full sample. The weights for the variables that were used to predict two outcomes are given herein, namely, death within 30 and 180 days after admission.[17]

Our goal was to create simple disease-specific scales that measured the risk of death that were similar in form to the APACHE II APS scale. We categorized and scaled all but five variables using our own clinical judgment or the APACHE II item scaling. For all variables, normal and missing values received a score of 0, and the maximum score was assigned to the people who were the most ill. For variables that defined the patients' medical and functional abnormalities, we assumed that patients did not have a condition unless it was mentioned in their medical record (Table 1).

In initial analyses, five variables (age, serum urea nitrogen level, temperature, systolic blood pressure, and coma) predicted in-hospital mortality especially well. We rescaled these variables, based on average mortality at different levels of the variable. We scored age, serum urea nitrogen level, and temperature by simply subtracting the mean value, across all patients with the disease, from the patient's value. Higher initial systolic blood pressure was a good sign for all diseases except cerebrovascular accident. The relative impact of confusion (altered mental or neurological status) and coma on death varied by disease, and we did a more clinically detailed scaling of coma for cerebrovascular accident, where it was overwhelmingly the strongest predictor of death.

Variable Selection and Weighting

We screened variables for use in the final regressions to predict mortality according to a rule that combined clinical judgment with statistical evidence. Variables were rated 1, 2, 3, and 4 by clinicians for their expected association with poor outcome. Using this rating, variables were kept for possible inclusion in the scale (1) if they were strong predictors of 30- or 180-day postadmission mortality (t statistic >2.5 in regressions that also included systolic blood pressure, age,

Table 1.—Scoring for Variables That Predict Death Within 30 or 180 Days After Admission*

Sickness-at-Admission Variable	Maximum Achievable Score†	Definition of Sickness-at-Admission Variable
APACHE II APS	53	Sum of 13 scales assessing patient's vital signs, blood studies, and mental status (CHF, AMI, PNE, HIP)‡
APACHE CHE§	1	Severe chronic failure of liver, cardiovascular, respiratory, or renal systems, or immunocompromised state = 1 (CHF, CVA, HIP)
Age (standardized)	. . .	(Age – mean sample age of 78 for CHF and PNE, 75 for AMI, 77 for CVA, 81 for HIP)/year
Systolic blood pressure	5	<90 mm Hg = 5, 90-99 mm Hg = 4, 110-119 mm Hg = 3, 120-139 mm Hg = 2, 140-159 mm Hg = 1, ≥160 mm Hg = 0 (5 diseases)
CHF by chest roentgenogram	3	Severe CHF by roentgenogram = 3 (PNE, CVA, HIP); moderate CHF by roentgenogram = 2 (CVA, HIP), 1 (PNE); mild CHF by roentgenogram = 1 (PNE, CVA, HIP); Pleural effusion by roentgenogram = 1 (CHF)
BUN (standardized)	. . .	(BUN – mean sample BUN of 10 mmol/L for CHF, 8 for AMI, 9.5 for PNE)/7 mmol/L (CHF, AMI, PNE)
APACHE II Coma Score‖	9	Coma = 9 (CHF, PNE); confusion or neurologic change = 5 (CHF, PNE, HIP)
CVA Coma Score	5	Posturing/no response to pain = 5; coma by both physician and nurse = 4; coma by physician or nurse = 3; neurologic change or unable to follow commands = 2; confusion = 1 (CVA)
From nursing home	1	Prior residence in skilled nursing facility, intermediate care facility, extended care facility, or nursing home with unspecified type = 1 (CHF, AMI, CVA, HIP)
Mean blood pressure‖	4	Mean blood pressure: ≥160 or ≤49 mm Hg = 4; 130-159 mm Hg = 3; 110-129 mm Hg or 50-69 mm Hg = 2 (CHF, CVA)
Heart rate‖	4	First heart rate value: ≥180 or ≤39 beats/min = 4; 140-179 or 40-54 beats/min = 3; 110-139 or 55-69 beats/min = 2 (AMI)
Serum sodium‖	4	First serum sodium value: ≥180 or ≤110 mmol/L = 4; 160-179 or 111-119 mmol/L = 3; 155-159 or 120-129 mmol/L = 2; 150-154 mmol/L = 1 (CHF)
Serum creatinine‖	4	First serum creatinine value: ≥300 μmol/L = 4; 180-300 μmol/L = 3; 133-168 or <53 μmol/L = 2 (HIP)
Oxygenation‖	4	Partial pressure of oxygen: Po_2 <55 or PAo_2-Pao_2 ≥ 500 = 4; Po_2 = 55-60 or PAo_2-Pao_2 = 350-499 = 3; PAo_2-Pao_2 = 200-349 = 2; Po_2 = 61-70 = 1 (AMI, HIP)
Hematocrit‖	4	Serum hematocrit: ≥0.60 or <0.20 = 4; 0.50-0.60 or 0.20-0.30 = 2; 0.46-0.50 = 1 (CHF)
White blood cell count‖	4	Serum white blood cell count: ≥40 or <1 × 10^9/L = 4; 20-39.9 or 1-2.9 × 10^9/L = 2; 15-19.9 × 10^9/L = 1 (CVA)
Digitalis toxic effects	4	Toxic digitalis level: >2.8 nmol/L = 4 (CHF)
Serum aspartate aminotransferase (AST)	4	Serum AST value: ≥80 U/L = 4 (CHF), 2 (AMI); 40-79 U/L = 2 (CHF), 0 (AMI)
Creatine kinase Score	4	Creatine kinase value: >3 × laboratory's highest normal value = 4; <3 but >1.3 × laboratory's highest normal value = 2 (AMI)
Serum albumin	4	Serum albumin value: <25 g/L = 4 (CVA, HIP), = 3 (PNE); 25-30 g/L = 2 (PNE, CVA, HIP); >30 g/L or not tested, but patient described as malnourished = 3 (HIP, PNE)
Male	1	Male = 1 (CHF, HIP)
Noncompliance	1	Noncompliance with medications or diet = 1 (CHF)
Use of ventricular depressant	1	Use of ventricular depressants (eg, verapamil, disopyramide phosphate, clonidine hydrochloride) = 1 (CHF)
Antiarrhythmic medications	1	Use of antiarrhythmic medication = 1 (CHF)
Difficulty with limbs	1	Difficulty with arm and/or leg function (CHF)
Prior hospitalization	1	Hospitalization within 6 mo = 1 (CHF, PNE)
Location of AMI	2	Location of AMI: anterior = 2; inferior or transmural but unspecified = 1; subendocardial = 0 (AMI)
Cardiac disease¶	8.7	Age >80 = 3.9; age 70-79 = 3.1; age <70 = 2.3; cardiomegaly by roentgenogram = 1.1; severe CHF by roentgenogram = 3.3; moderate CHF by roentgenogram = 2; mild CHF by roentgenogram or lungs not clear on examination = 1; prior angina = 0.4 (AMI)
Prior CHF	1	History of CHF or use of digitalis at time of admission = 1 (AMI)
Prior valve disease	1	History of mitral, tricuspid, aortic, pulmonic valve disease, or valvular surgery = 1 (AMI, CVA)

(Continued on p 1964.)

Table 1.—Scoring for Variables That Predict Death Within 30 or 180 Days After Admission* (cont)

Sickness-at-Admission Variable	Maximum Achievable Score†	Definition of Sickness-at-Admission Variable
Prior diabetes	1	History of diabetes = 1 (AMI, CVA)
Intubation	1	Oropharyngeal or nasopharyngeal intubation = 1 (AMI)
Body System Score	12	Count of No. of body systems (eg, cardiovascular, respiratory) involved with acute or chronic disease (AMI)
Infarct type	3	Hemorrhagic infarct = 3; lacunar infarct = −1 (CVA)
Shortness of breath	1	Dyspnea noted by physician and/or nurse = 1 (PNE, CVA)
Arrest in emergency department	1	Cardiac and/or respiratory arrest or intubation in emergency department = 1 (CVA)
Recent CHF	1	CHF diagnosed as active within 1 wk of admission = 1 (CVA)
Malnutrition	1	Noted by physician at admission = 1 (CVA)
Nonambulatory on day 1	1	Noted by physician or nurse as nonambulatory on day 1 (CVA)
Respiratory distress	1	Noted by physician (PNE)
Septic complications	1	Empyema, lung abscess, pericarditis, or meningitis present at admission (PNE)
Prior respiratory failure	1	Noted by physician (PNE)
Thoracic disease	1	Neuromuscular disease (eg, amyotrophic lateral sclerosis, myasthenia gravis) = 1 (PNE)
Prior cancer score	3	Poor-prognosis cancer or cancer treated by chemotherapy or radiation = 3; cancer treated with hormone therapy = 2; cancer with no evidence of activity = 1 (PNE)
New lung cancer	1	New cancer diagnosis made during hospitalization = 1 (PNE)
Home oxygen	1	Use of home oxygen prior to admission = 1 (PNE)
Nonambulatory preadmission	1	Noted by physician or nurse as nonambulatory prior to admission (PNE)
RAND Comorbidity Scale	45	Weighted sum of chronic comorbid conditions present at admission (PNE, HIP)#
Temperature (standardized)	. . .	(Temperature–37.6)°C (HIP)
Pneumonia by chest roentgenogram	1	Chest roentgenogram report: bilateral pneumonia without CHF = 1; unilateral pneumonia or infiltrate = 0.5 (HIP)
Prior renal failure	1	History of renal transplant or long-term dialysis, and creatinine value on admission ≥2.0 mg/dL (HIP)
Prior CVA	1	CVA prior to admission = 1 (HIP)
Prior COPD	1	COPD, emphysema, asthma, or chronic bronchitis prior to admission (HIP)

*APACHE indicates *A*cute *P*hysiology *a*nd *C*hronic *H*ealth *E*valuation; APS, acute physiological score; CHF, congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; HIP, hip fracture; CVA, cerebrovascular accident; BUN, serum urea nitrogen; and COPD, chronic obstructive pulmonary disease. Values are given only for variables included in the scales that predict mortality 30 and/or 180 days after admission. Laboratory values are given in Système International (SI) units.

†The maximum achievable score is the difference between the healthiest and sickest categories in all instances where the healthiest category is scored 0. This is the case for all variables except "cardiac disease score" and "infarct type," BUN level, age, and temperature.

‡The variable was used in the final sickness-at-admission score for the diseases in parentheses.

§CHE indicates chronic health evaluation from the APACHE CHE measure.

‖A component of the APACHE APS score. We use the standard APACHE scores.

¶Age, cardiomegaly, pulmonary congestion, and history of ischemia as defined by the Norris Coronary Prognostic Index.

#In the sum (with weights in parentheses) are 16 measures of prior conditions: cancer (3), cirrhosis (2), diabetes (1), cerebrovascular accident (2), chronic renal failure (3), valvular disease or angina or myocardial infarction or heart surgery (2), congestive heart failure (1), arrhythmias (2), swallowing disorder (eg, aspiration, dysphagia) (2), use of nasogastric tube (3), hospitalization in the last month (2), thoracic or abdominal surgery in the last month (2), disease of the thorax (3), multiple myeloma (2), splenectomy (2), and dementia (2), and five measures of current problems: smoking (2), alcoholism (2), morbid obesity (2), hypoalbuminemia or malnourishment (3), and immunocompromised state (2).

and the APACHE II APS scale); (2) if they were medium predictors (t statistic >2) with a clinical score of 1 or 2; or (3) if the t statistic was greater than 1.5 and the clinicians expected a strong relationship with death (a clinical score of 3 or 4).

Finally, we used logistic regression to assign weights to the variables. The final weights were rounded to the nearest integer multiple of the APACHE II APS weights (ie, logistic regression coefficients were divided by the coefficient for APACHE II APS and rounded). Insignificant ($P>.05$) variables, and six measures of sickness implausibly associated with decreased mortality, were dropped. To estimate conservatively the predictive ability of these sickness-at-admission variables, we divided the sample into 10 equal parts at random. The R^2 values we report are the averages of the squares of the simple correlations between outcomes in each random tenth of the sample and predictions based on the other 90% of patients.[9]

Changes in Laboratory Ordering and Reporting

From 1981 to 1986, we observed substantial increases in the number of ordered tests for arterial blood gas levels, and lesser increases in the ordering of other tests. We were concerned that these differences might lead to artifactual increases in sickness at admission.

To investigate this, we divided the sickness items into three categories: recording-sensitive (Table 2), laboratory, and recording-insensitive. For analyses of changes over time we assumed, conservatively, that the change in the recording-insensitive part of the scale (ie, measurement of vital signs, sex, and age) represented the true change in sickness over time. We reduced the changes that we observed in laboratory abnormalities using missing laboratory value indicators, and we reduced the post-PPS weights of recording-sensitive items uniformly to make the changes in these categories proportional to the changes in vital signs and age. To the extent that the increase in tests and reported comorbidity is due to the increased illness in patients, rather than to changed medical practice, these adjustments will lead us to underestimate increases in sickness at admission following the introduction of prospective payment.

RESULTS

Sickness-at-Admission Scales

Sickness-at-admission variables and their weights that were used to predict 30- or 180-day postadmission mortality are given in Table 2. The weights reflect relative importance, eg, in patients with pneumonia, the presence of respiratory distress (with weight 6) is equivalent in its effect on mortality to the patients' being 6/0.45 or 13.3 years older (each year of age has weight 0.45).

The models based on the bootstrap and validation data sets were similar to those presented in Table 2. The choice of variables, especially rare variables, and the weights varied from replicate to replicate, but the average pairwise correlation of predicted death probabilities from the different models was .94.

The numbers at the bottom of Table 2 are used to convert the sickness-at-admission score for a patient to the predicted probability of death within 30 days, assuming average process of care in the 1985 to 1986 sample. The multiplier is the change in the log odds of dying of each one-point change in the sickness-at-admission score, and the intercept is the log odds of dying for people with no abnor-

Table 2.—Weights for Sickness-at-Admission Variables That Predict 30-Day Postadmission Mortality*

Variable	Observed Range	Weights†				
		CHF	AMI	PNE	CVA	HIP
APACHE II APS‡	33	0.5	1**	1*		1**
Age (standardized)	40	0.2	0.4*	0.45*	0.25*	0.15*
Blood pressure score	5	4***	3.5***	2*	2*	1*
CHF by chest roentgenogram	3	4*		3	3	3*
BUN (standardized)	9	3**	2	3*		
Coma Score‡	9	0.5		0.77		−0.8*
CVA Coma Score‡	5				5**	
Mean blood pressure	4				1	
Heart rate	4		−1			
Serum sodium	4	2				
Oxygenation	4					2*
Hematocrit	4	−0.5				
White blood cell count	4				4*	
Digitalis toxic effects	4	1.5				
Serum aspartate aminotransferase	4	1.5*	3.5**			
Creatinine kinase Score	4		−1.5*			
Infarct type	4				3**	
Serum albumin	3			2		1*
Location of myocardial infarction	2		6**			
Intubation‡	1		9*			
Body system count‡	6		2*			
Shortness of breath‡	1			7*	5*	
Arrest in emergency department‡	1				9	
Recent CHF‡	1				12*	
Prior diabetes‡	1				4*	
From nursing home‡	1				3	
Respiratory distress‡	1			6*		
Septic complications‡	1			20*		
Prior respiratory failure‡	1			10		
Thoracic disease‡	1			15		
Recently hospitalized‡	1			4		
Nonambulatory‡	1			12*		
Temperature (standardized)	10			−3*		
Male	1					3*
Pneumonia by chest roentgenogram	1					6
Prior renal failure‡	1					18*
Prior CVA‡	1					3*
Prior COPD‡	1					2
1985-1986 multiplier		0.102	0.082	0.082	0.115	0.112
1985-1986 intercept		−3.43	−3.11	−4.22	−3.96	−3.99
1985-1986 30-day postadmission mortality		0.131	0.248	0.170	0.199	0.048

*APACHE indicates *A*cute *P*hysiology and *C*hronic *H*ealth *E*valuation; APS, acute physiological score; CHF, congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; HIP, hip fracture; COPD, chronic obstructive pulmonary disease; and BUN, serum urea nitrogen. Significance is shown by asterisks beside numerals: no asterisk indicates $2<|t|<3$; one asterisk, $3<|t|<5$; two asterisks, $5<|t|<7$; and three asterisks, $|t|>7$.

†Logistic regression coefficients were rounded to simple multiples of the APACHE coefficient.

‡Recording-sensitive items (in APACHE II APS only the Coma Score is recording-sensitive).

malities (Table 3).

The predictions of death for particular diseases are based on those variables for which numbers are provided in Table 2. In particular, we tested 64, 61, 82, 83, and 74 variables, respectively, for congestive heart failure, acute myocardial infarction, cerebrovascular accident, pneumonia, and hip fracture, and concluded with the inclusion of 18, 20, 12, 25, and 22 variables, respectively, in the final disease-specific measures (counting each of the 13 variables in APACHE II APS separately). All variables in the scales are significant predictors ($P<.05$). The APACHE II APS score, rescored low blood pressure, serum urea nitrogen level, coma, age, and congestive heart failure on chest roentgenogram were important predictors for most diseases. Other variables, such as elevated aspartate aminotransferase level and infarct location for patients who had a myocardial infarction, or cerebral infarct type for patients who had a cerebrovascular accident, were important predictors for specific diseases.

The 180-day mortality scales include the 30-day mortality scales and additional variables (Table 4). The 30-day mortality scales are the most important part of the 180-day mortality scales, both because 180-day mortality includes 30-day mortality and because many variables that influence short-term mortality have prognostic value over the longer term as well. Age, prior nursing home residence, and chronic conditions are important predictors of longer-term mortality.

Predictive Power of the Sickness Scales

Table 5 presents conservative estimates of the predictive power of the RAND sickness-at-admission measures, and compares them with the predictive performance of other measures. In particular, we examined the APACHE II APS scale, which uses the same scoring and weighting of its items for all diseases; the APACHE II variables with rescaled blood pressure, serum urea nitrogen level, coma, and disease-specific weights (lines 3 and 4 of Table 5); and the scales used in the Medicare Mortality Predictor System (MMPS).[9] Age-sex adjustment does not explain much of the variance in 30-day mortality in patients aged 65 years and older. The addition of the generic APACHE II APS adds to explanatory power, but the power of the APACHE II variables is increased if they are reweighted specifically for each disease. The big increases in prediction for acute myocardial infarction and congestive heart failure are mainly due to the greatly increased weight on systolic blood pressure. The complete RAND measures offer further improvement. The RAND measures explain about one quarter of the variance in 30-day mortality at the patient level for patients who have a myocardial infarction, pneumonia, or cerebrovascular accident.

Table 6 shows the 30- and 180-day postadmission mortality of those patients who were rated healthiest and sickest by the RAND sickness-at-admission scales. The differences between the healthiest and sickest quartiles are clinically meaningful. For example, of those patients whose admission characteristics put them in the healthiest 25% of patients who had a cerebrovascular accident, 3% died in the 30 days that followed admission, whereas 55% of patients in the sickest quartile died (a relative risk of 20; 95% confidence limits, 13 to 32). The 180-day prognosis contrasts are almost as large.

Comparisons Between Pre- and Post-PPS

Sickness at admission increased significantly following the introduction of pro-

Table 3.—Calculation of Predicted Death Probabilities for a Hypothetical Patient With Pneumonia*

Assumed Characteristics	Item Score (From Table 1)	Weight (From Table 2)	Product
Heart rate, blood pressure, potassium and white blood cell abnormalities for an APACHE II APS score of 10	10	1.0	10.0
88 years old (88 − 78 [mean age] = 10)	10	0.45	4.5
Systolic blood pressure 80 mm Hg	5	2.0	10.0
BUN and temperature same as sample mean	0	...	0
Recently hospitalized	1	4.0	4.0
No data about preadmission ambulation	0	...	0.0
All other characteristics normal	...	...	0.0
(30-Day Mortality) Sickness-at-Admission Score	...	...	**28.5**

*APACHE indicates Acute Physiology and Chronic Health Evaluation; APS, acute physiology score; and BUN, serum urea nitrogen. We use the 1985 to 1986 multiplier and intercept of 0.082 and −4.22 (from Table 2) to estimate that patient A has a 13.2% probability of dying within 30 days postadmission, assuming average care in the 1985 to 1986 period as follows: (28.5 × 0.082) − 4.22 = −1.88, the logit of dying, and exp (−1.88)/[1 + exp(−1.88)] = 13.2%.

Table 4.—Weights for Sickness-at-Admission Variables That Predict 180-Day Postadmission Mortality*

Variable	Observed Range	Weights†				
		CHF	AMI	PNE	CVA	HIP
30-d scale	98	1***	1***	1***	1***	1***
Age (standardized)	40			0.15	0.2	0.35*
APACHE II CHE	1	4			7	10*
CHF by chest roentgenogram	3				2	
BUN (standardized)	9	4**	4			
Male	1	2*				3
From nursing home	1	4	6		6	6*
Mean blood pressure	4	−1*				
Serum creatinine	4					2*
Oxygenation	4		1			
Coma score	5					0.8
Cardiac disease‡	6.4		3**			
Serum albumin	3				4*	
Infarct type	4				−2	
Prior hospitalization	1	3*				
Difficulty with limbs	1	4				
Noncompliance	1	−5				
Use of ventricular depressants	1	−5				
Antiarrhythmic medications	1	4				
Prior CHF	1		6*			
Prior valve disease	1		8		6	
Prior diabetes	1		6			
Malnutrition	1				6	
Nonambulatory on day 1	1				7	
Nonambulatory (prior to admission)	1			6		
Prior cancer score	3			4*		
New lung cancer	1			24*		
Home oxygen	1			6		
RAND Comorbidity Scale	16			1*		1*
Prior renal failure	1					−18
Pneumonia by chest roentgenogram	1					−6
1985-1986 multiplier		0.086	0.065	0.069	0.089	0.079
1985-1986 intercept		−2.10	−3.12	−3.28	−3.27	−3.12
1985-1986 180-day postadmission mortality		0.320	0.359	0.313	0.342	0.152

*APACHE indicates Acute Physiology and Chronic Health Evaluation; CHE, chronic health evaluation; CHF, congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; HIP, hip fracture; and BUN, serum urea nitrogen. Significance is shown by asterisks beside numerals: no asterisk indicates $2<|t|<3$; one asterisk, $3<|t|<5$; two asterisks, $5<|t|<7$; and three asterisks, $|t|>7$.

†Logistic regression coefficients were rounded to simple multiples of the 30-day score coefficient.

‡Age, cardiomegaly, pulmonary congestion, and history of ischemia.

spective payment (Table 7). For all five diseases combined, the 30-day mortality to be expected from increased sickness at admission was 1.0% higher in the period from 1985 to 1986 (16.4% vs 15.4%, $P<.01$; 95% confidence interval [CI] for difference, 0.3% to 1.7%), and the expected 180-day mortality was 1.6% higher (30.1% vs 28.5%, $P<.001$; 95% CI for difference, 0.8% to 2.4%). Patients with pneumonia had the greatest increase over time in expected mortality from increased sickness at admission. If we had not adjusted for changes in laboratory ordering and recording, the increases in sickness post-PPS in Table 7 would have been an additional 0.6 percentage points of expected mortality higher than the indicated values (eg, 1.6 percentage points of expected 30-day mortality and 2.2 percentage points of expected 180-day mortality).

COMMENT

Apart from the MMPS,[9] previous scales to measure sickness at admission (1) have most often been applied to patients who were hospitalized in the intensive care unit, (2) are based on data collected at any time during the admission, (3) are not easily generalizable, or (4) are based on variables and algorithms that are not in the public domain. The sickness-at-admission scales presented herein were developed from a large nationally representative sample of more than 14 000 patients with one of five diseases at 297 hospitals. They are publicly available, and could be used at either the individual patient level or the hospital level. For example, a physician who treats a patient with one of our five conditions could compute our sickness-at-admission value to find whether the predicted death probability for this patient was 5% or 50%. Hospitals that differ in their proportion of healthy and ill patients will have different outcomes that are independent of the care they provide; our scales could help to remove this confounding factor in hospital comparisons.

Our results have implications for those who want to use APACHE APS or MMPS. The APACHE II variables were important in predicting death for all patients, but blood pressure and renal failure needed to be rescaled for our cohort of elderly patients. In contrast to APACHE's results on patients who were hospitalized in intensive care units,[3] high values of blood pressure were more favorable than normal values, especially for patients with cardiac diseases, and reweighting for each disease led to better predictions. Serum urea nitrogen level was important even after accounting for serum creatinine values. In addition, we identified additional laboratory findings, eg, serum

Table 5.—Predictive Power of the Sickness-at-Admission Scales for 30-Day Postadmission Mortality: Percent of Variance Explained*

Predictors	R^2, %				
	CHF	AMI	PNE	CVA	HIP
Age, sex	1	2	3	2	1
Age, sex, APACHE II APS	3	12	18	17	3
Reweighted APACHE II APS†	10	18	21	26	3
Reweighted APACHE II APS and Body System Count†	10	18	23	26	5
MMPS†‡	15	14	18	25	...
RAND 30-day scale†	12	22	26	30	6

*APACHE indicates Acute Physiology and Chronic Health Evaluation; APS, acute physiological score; CHF, congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.
†R^2 computed in conservative way (see text).
‡Medicare Mortality Predictor System. Values computed on different but comparable datasets.

Table 6.—Post admission Mortality by Sickness-at-Admission Quartile for Five Diseases*

Sickness-at-Admission Quartile†	Postadmission Mortality Rates, %					
	CHF	AMI	PNE	CVA	HIP	Five Diseases
	30 Days					
Healthiest quartile	5	9	1	3	1	2
Middle quartiles	9	17	9	13	4	10
Sickest quartile	32	55	45	55	10	43
Overall	14	24	16	21	5	16
	180 Days					
Healthiest quartile	11	11	4	9	3	7
Middle quartiles	29	28	22	29	14	24
Sickest quartile	61	70	67	72	34	64
Overall	32	34	29	35	16	29

*CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.
†Sickness-at-admission quartiles were determined by rank ordering patients according to their sickness-at-admission score.

Table 7.—Sickness at Time of Hospital Admission: PRE-PPS (1981-1982) vs POST-PPS (1985-1986)*

Sickness at Admission	Expected Mortality, %†					
	CHF	AMI	PNE	CVA	HIP	Five Diseases
30-day scale						
Pre-PPS	13.6	23.5	14.2	20.6	4.6	15.4
Post-PPS	13.9	24.8	17.1	20.6	5.2	16.4
Difference (post-PPS – pre-PPS)	**0.3**	**1.3**	**2.9*****	**0.0**	**0.6****	**1.0****
180-day scale						
Pre-PPS	32.2	33.2	26.2	34.9	15.6	28.5
Post-PPS	32.8	35.2	30.7	34.7	16.8	30.1
Difference (post-PPS – pre-PPS)	**0.6**	**2.1***	**4.5*****	**−0.2**	**1.2***	**1.6*****
Sample sizes	2591	2588	2536	2565	2541	12 821

*PPS indicates prospective payment system; CHF, Congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA; cerebrovascular accident; and HIP, hip fracture. Significance of pre-PPS and post-PPS differences is shown by asterisks beside numerals: one asterisk indicates $P<.05$; two asterisks, $P<.01$; and three asterisks, $P<.001$.
†Expected mortality based on sickness at admission, assuming average in-hospital processes of care over the entire study.

aspartate aminotransferase and serum albumin levels, that were important predictors for particular diseases.

Our scales combine the acute data that are the focus of the MMPS with an inventory of chronic morbid and comorbid disease markers. By using additional data, our predictive results are usually better than those of the MMPS. These results confirm earlier findings that chronic comorbidity is important for predicting longer-term outcomes.[18,19] Considerations in study design kept us from adjusting for the assignment of a "do not resuscitate" order or for metastatic cancer (ie, adjustment for a do not resuscitate order might have confounded the PPS effect and we excluded patients with metastatic cancer), which in the MMPS were strong predictors. We could have raised predictive power by including these variables or patients. If even more explanatory power is desired, then a data source other than the medical record will be required (eg, data collected prospectively from patients and/or physicians).

The decision about which sickness-at-admission measures to use depends on the level of clinical data an individual is willing to collect. Whether it is worth the cost to collect our additional clinical variables routinely is an important area for further study. Based on work reported herein, to strengthen the mortality models that are used for the annual hospital mortality reports from the Health Care Financing Administration[20] we would recommend first adding data on initial systolic blood pressure, and then on a few additional laboratory variables.

The number of variables that are available from the medical record and are important in measuring sickness at admission is limited. We began with an average of 73 sickness variables per disease, but our final models for sickness measures that predict 30-day postadmission mortality contained an average of 19. Many variables that are strongly associated with death when considered by themselves are not part of our final scales. Usually, such variables are pushed aside by a better measure of the same problem. For example, current laboratory measures of renal failure overshadow a history of renal failure, except for patients with hip fractures.

In this study, we dealt with the problem of differences in recording and laboratory ordering styles among hospitals and over time. This problem must be addressed in any comparative study where such differences are expected. Each comparison requires a study of recording and laboratory ordering practices and appropriate adjustment. We have provided one (conservative) way to do this.

We have shown that patients, especially those with pneumonia, were more ill at admission in the period that followed implementation of prospective payment. This increase in illness must be taken into account in evaluating the changes in posthospital death rates over time. Because hospitals, following introduction of prospective payment, had financial incentives to hospitalize patients who were not as ill, the apparent increase in sickness is initially puzzling. We have adjusted for increases in the ordering of laboratory tests and recording of comorbidity, so they cannot explain the rise. Vital signs at admission are significantly worse for patients who have a myocardial infarction, pneumonia, and hip fracture, indicating that these patients unequivocally were more ill at admission.

So why are hospitalized patients more ill on average than they used to be? Better paramedical services may keep more

ill people alive to be hospitalized, and financial incentives to increase admission of patients who are not as ill may be less important than are activities of professional review organizations, increased external review of appropriateness of hospitalizations, or shifts from inpatient to outpatient settings for treatment. Hospital admissions have fallen,[21] despite the financial incentives, and these other trends may have prevented the hospitalization of many patients who were not as ill, leaving a population with increased illness in the hospitals.

We have developed sickness-at-admission measures for five diseases that account for one third of all hospital deaths in the Medicare population. These measures are disease-specific, in the public domain, and explain about one fourth of the variance in death at the patient level following hospitalization for acute myocardial infarction, pneumonia, or cerebrovascular accident. Our sickness measures demonstrate a more than 10-fold variation in actual death rates between the upper and lower quartiles of patients with each disease. Finally, we have shown that increases in sickness at admission accompanied the unexpected drop in admissions following the introduction of prospective payment.

This study was supported by cooperative agreement 18-C-98853/9-03 issued by the Health Care Financing Administration of the US Department of Health and Human Services, Baltimore, Md. The entire RAND cost (3.9 million) of this project was funded with federal money.

We gratefully acknowledge the collaborative effort of the five professional review organizations that enabled this work to be completed. Participants from these organizations included medical directors, physician consultants and reviewers, project directors, and review coordinators from each of the five study states. In particular, we appreciate the keen clinical insight of the physician specialists with whom we consulted throughout this study. We also acknowledge the many contributions of Harry Savitt, PhD, project officer from the Office of Research and Demonstrations of the Health Care Financing Administration, US Department of Health and Human Services, whose administrative skills, astute commentary, and continuing support helped us immeasurably. Maureen Carney, MS, Carole Chew, RRA, MPH, Patricia A. Damiano, Linda Ferguson, Caren Kamberg, MSPH, Nancy J. Palmer, and Carol Roth, RN, MPH, were key contributors to project development and implementation. Andrea Steiner, MS, gave valuable editorial advice. Our policy advisory board (John C. Beck, MD, Barbara J. Burns, PhD, Monroe T. Gilmour, MD, Paul F. Griner, MD, Charlene Harrington, RN, PhD, T. Reginald Harris, MD, Rosalie Kane, DSW, Shirley Kellie, MD, Judith R. Lave, PhD, Charles E. Lewis, MD, Joseph Martin, Francis D. Moore, Jr, MD, Richard N. Pierson, Jr, MD, and James F. Rodgers, PhD) provided astute advice and guidance. We acknowledge the important contributions of the 297 hospitals whose medical records we reviewed. Without the efforts of these individuals and institutions, this evaluation could not have been successfully completed. Finally, we recognize and thank Martha B. Cooper, without whose secretarial skills this article would not have been produced.

References

1. 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.
2. Kahn KL, Brook RH, Draper D, et al. Interpreting hospital mortality data. *JAMA.* 1988;260:3625-3628.
3. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: a severity of disease classification system. *Crit Care Med.* 1985;13:818-829.
4. Kruse JA, Thill-Baharozian MC, Carlson RW. Comparison of clinical assessment with APACHE II for predicting mortality risk in patients admitted to a medical intensive care unit. *JAMA.* 1988;260:1739-1742.
5. Lemeshow W, Teres D, Pastides H, et al. A method for predicting survival and mortality of ICU patients using objectively derived weights. *Crit Care Med.* 1985;13:519-525.
6. Killip T, Kimball JT. Treatment of myocardial infarction in a coronary care unit. *Am J Cardiol.* 1967;20:457-461.
7. Norris RM, Brandt PWT, Caughey DE, et al. A new coronary prognostic index. *Lancet.* 1969;1:274-278.
8. Goldman L, Caldera DL, Nussbaum SR, et al. Multifactorial index of cardiac risk in noncardiac surgical procedures. *N Engl J Med.* 1977;297:845-850.
9. Daley J, Jencks SF, Draper D, et al. Predicting hospital-associated mortality for Medicare patients. *JAMA.* 1988;260:3617-3624.
10. 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 Corp; 1988. Publication N-2798-HCFA.
11. 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 Corp; 1988. Publication N-2799-HCFA.
12. Roth CP, Kahn KL, Sherwood MJ, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Pneumonia.* Santa Monica, Calif: RAND Corp; 1988. Publication N-2801-HCFA.
13. Rubenstein LV, Kosecoff J, Kahn KL. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Cerebrovascular Accident.* Santa Monica, Calif: RAND Corp; 1988. Publication N-2802-HCFA.
14. 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 Corp; 1988. Publication N-2800-HCFA.
15. Kahn KL, Draper D, Keeler EB, et al. *The Effects of DRG-Based Prospective Payment on Quality of Care for Hospitalized Medicare Patients.* Santa Monica, Calif: RAND Corp; 1990. Publication R-3931-HCFA.
16. Altman DG, Anderson PK. Bootstrap investigation of the stability of a Cox regression model. *Stat Med.* 1989;8:771-783.
17. Kahn KL, Keeler EK, Sherwood MS, et al. Comparing outcomes of care pre- and post-implementation of the DRG-based prospective payment system. *JAMA.* 1990;264:1984-1988.
18. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis.* 1987;40:373-383.
19. Clemens JD, Feinstein AR, Holabird N, Cartwright S. A new clinical—anatomic staging system for evaluating prognosis and treatment of prostatic cancer. *J Chronic Dis.* 1986;39:913-928.
20. *Medicare Hospital Mortality Information.* Washington, DC: Health Care Financing Administration; 1987. Publication No. 00642.
21. Guterman S, Eggers PW, Riley G, et al. The first three years of Medicare prospective payment: an overview. *Health Care Financ Rev.* 1988;9:67-77.

Measuring Quality of Care With Explicit Process Criteria Before and After Implementation of the DRG-Based Prospective Payment System

Katherine L. Kahn, MD; William H. Rogers, PhD; Lisa V. Rubenstein, MD, MSPH; Marjorie J. Sherwood, MD; Ellen J. Reinisch, MS; Emmett B. Keeler, PhD; David Draper, PhD; Jacqueline Kosecoff, PhD; Robert H. Brook, MD, ScD

We developed explicit process criteria and scales for Medicare patients hospitalized with congestive heart failure, myocardial infarction, pneumonia, cerebrovascular accident, and hip fracture. We applied the process scales to a nationally representative sample of 14 012 patients hospitalized before and after the implementation of the diagnosis related group–based prospective payment system. For the four medical diseases, a better process of care resulted in lower mortality rates 30 days after admission. Patients in the upper quartile of process scores had a 30-day mortality rate 5% lower than that of patients in the lower quartile. The process of care improved after the introduction of the prospective payment system; eg, better nursing care after the introduction of the prospective payment system was associated with an expected decrease in 30-day mortality rates in pneumonia patients of 0.8 percentage points, and better physician cognitive performance was associated with an expected decrease in 30-day mortality rates of 0.4 percentage points. Overall, process improvements across all four medical conditions were associated with a 1 percentage point reduction in 30-day mortality rates after the introduction of the prospective payment system.

(*JAMA*. 1990;264:1969-1973)

PROCESSES of care—what we do to patients—have been considered an essential component of quality of care measurement for over 50 years.[1-6] Even if outcomes of care—what happens to patients—are the most meaningful measures of quality to the patient, we will be unable to develop clinical methods to improve outcomes unless we understand how processes and outcomes are related. Assessing quality of care by process also provides some measurement advantages over studying outcomes, because not all patients who experience a poor process of care suffer a poor outcome.

The purpose of this article is twofold. First, we report on the development of a set of validated process criteria for elderly patients admitted to the hospital with one of five conditions. By validated we mean that process predicts outcome. Second, we apply the validated process criteria to patients treated before and after the implementation of the prospective payment system (PPS) to determine whether the PPS has been associated with changes in the processes of care.

METHODS

We based our analysis on the sample described in more detail elsewhere in this series.[7]

Developing Process Criteria

We used literature review and consultation with experts to develop a set of process measures for which better process was likely to make a difference in patient outcome. These measures were then presented to disease-specific panels consisting of five to 12 physicians, who were selected by our collaborators, the professional review organizations. Each panel reviewed the suggested criteria to decide whether they believed that data to assess these criteria were reliably recorded in the medical record and whether the criteria made clinical sense. Process criteria based on data whose recording was likely to vary by year, state, or hospital type were excluded. We developed disease-specific abstraction forms[8-12] to collect data on approximately 100 process criteria for each disease.

Scoring Process Criteria

In scoring process criteria, we first applied the criteria only to patients who were likely to benefit from their use. Using this kind of conditional logic, many criteria were applicable to all patients, some to just a few. For example, *if* a patient with congestive heart failure was considered to be severely ill, *then* the intensive care unit should be used. Second, we used clinical judgment to assign scores (points) to each process criterion based on how likely a patient was to benefit from it. For example, use of the intensive care unit for very sick patients was assigned seven points, whereas use of the intensive care unit for moderately sick patients was assigned three points. Third, the process

From the Health Program of the RAND Corp, Santa Monica, Calif (Drs Kahn, Rogers, Rubenstein, Sherwood, Keeler, Draper, and Brook and Ms Reinisch); the Departments of Medicine (Drs Kahn, Rubenstein, Kosecoff, and Brook) and Health Services (Drs Kosecoff and Brook), UCLA; and Value Health Sciences Inc, Santa Monica (Dr Kosecoff).

The opinions, conclusions, and proposals in the text are those of the authors alone and do not necessarily represent the views of the RAND Corp or UCLA.

Reprint requests to the RAND Corp, 1700 Main St, PO Box 2138, Santa Monica, CA 90406-2138 (Dr Kahn).

Reprinted from *JAMA, The Journal of the American Medical Association*, October 17, 1990, Vol. 264, No. 15, pp. 1969–1973, © 1990, American Medical Association. Reprinted by permission.

Table 1.—Examples of Process Criteria and Performance Levels Before and After Introduction of the PPS*

Criteria	Disease	Patients to Whom Criteria Were Applicable, %		Patients to Whom Criteria Were Applicable Who Met Process Criteria, %	
		Before PPS	After PPS	Before PPS	After PPS
Physician Cognitive Scale					
Within the initial 2 days of hospitalization the physician should document each of the following in the medical record as noted or not noted					
Past surgery	Congestive heart failure	100	100	61	66†
Lung examination on day 2	Congestive heart failure	100	100	58	71†
Alcoholism or smoking habits	Acute myocardial infarction	100	100	61	64
Jugular veins	Acute myocardial infarction	100	100	61	68†
Tobacco use or nonuse	Pneumonia	100	100	47	52†
Lower-extremity edema	Pneumonia	100	100	68	75†
Previous cerebrovascular accident	Cerebrovascular accident	100	100	48	53†
Gag reflex	Cerebrovascular accident	100	100	35	38
Mental status	Hip fracture	100	100	68	70
Pedal or leg pulse	Hip fracture	100	100	62	67†
Nurse Cognitive Scale					
On day 2 of the hospitalization at least three blood pressure readings should be noted					
>3 blood pressure readings noted	Congestive heart failure	100	100	78	84†
>3 blood pressure readings noted	Pneumonia	100	100	69	79†
>3 blood pressure readings noted	Cerebrovascular accident	100	100	79	86†
Technical Diagnostic Scale					
Within the initial 2 days of hospitalization an electrocardiogram should be obtained					
Electrocardiogram obtained	Congestive heart failure	100	100	87	91†
Electrocardiogram obtained	Cerebrovascular accident	100	100	82	86†
Electrocardiogram obtained	Hip fracture	100	100	90	93†
Within the initial 2 days of hospitalization a serum potassium determination should be performed					
Serum potassium level determined	Congestive heart failure	100	100	93	97†
Serum potassium level determined	Cerebrovascular accident	100	100	88	94†
Serum potassium level determined	Hip fracture	100	100	89	94†
Technical Therapeutic Scale					
If po_2<60 mm Hg, use oxygen therapy or intubate					
Oxygen therapy or intubation done	Congestive heart failure	16	20	87	93†
Oxygen therapy or intubation done	Pneumonia	23	31	83	90†
Begin antibiotic therapy for patients with pneumonia in a timely manner					
Within 4 hours of admission for nonimmunocompromised patients	Pneumonia	91	88	28	32†
Within 2 hours of admission for nonimmunocompromised patients	Pneumonia	9	12	3	4
Monitoring With Intensive Care and Telemetry Scale					
For patients who are moderately sick‡ use the intensive care unit; telemetry is not sufficient but is preferable to no cardiac monitoring					
Used intensive care unit on day 1	Congestive heart failure	16	16	43	46
Used telemetry on day 1	Congestive heart failure	16	16	8	24†
Used intensive care unit on day 2	Congestive heart failure	6	7	49	41
Used telemetry on day 2	Congestive heart failure	6	7	13	31†
Used intensive care unit on day 1	Pneumonia	17	19	7	9
Used telemetry on day 1	Pneumonia	17	19	4	10†
Used intensive care unit on day 2	Pneumonia	8	8	31	35
Used telemetry on day 2	Pneumonia	8	8	4	12†
For patients who are very sick‡ use the intensive care unit; telemetry is not sufficient but is preferable to no cardiac monitoring					
Used intensive care unit on day 1	Congestive heart failure	5	6	61	71
Used telemetry on day 1	Congestive heart failure	5	6	12	17
Used intensive care unit on day 2	Pneumonia	9	11	42	50
Used telemetry on day 2	Pneumonia	9	11	5	7

*PPS indicates prospective payment system.
†$P<.05$ compared with before PPS.
‡Moderately sick was defined as a score of 5 or 6 and very sick as a score ≥7 on each hospital day, with points assigned as follows: chest pain, 1 point; shortness of breath, 1 point; confusion, 2 points; heart rate ≥130 beats per minute, 2 points; respiratory rate ≥30/min, 2 points; and diastolic blood pressure ≥105 mm Hg and systolic blood pressure <90 mm Hg, 3 points.

scores accounted for the use of different interventions. Very sick patients received seven of seven points for use of the intensive care unit, three of seven points for use of telemetry, and no points for no cardiac monitoring.

Process Scales

Using clinical judgment we grouped process criteria according to what concept we thought they measured and then tested our groupings by comparing them with those suggested by a Likert scaling model.[13] Use of these methods yielded five process subscales and one overall process scale: physician cognitive, nurse cognitive, technical diagnostic, technical therapeutic, monitoring

with intensive care or telemetry, and overall process.

The physician diagnostic cognitive scale measures the physician's performance as a gatherer of data about the patient's medical history and current symptoms and the performance of physical examinations during the hospital stay. The nurse diagnostic cognitive scale measures the nurse's performance as a gatherer of data about the patient's functional status, current symptoms, and vital signs. The technical diagnostic process scale measures use of diagnostic tests (eg, venous laboratory studies, arterial blood gas tests, roentgenograms, and electrocardiograms) that are indicated given the patient's daily burden of illness. The technical therapeutic process scale measures use of treatments (eg, medication, surgery, and physical therapy) that are indicated given the patient's daily sickness level. The intensive care or telemetry monitoring scale evaluates the monitoring of patients as a function of their level of illness. Whereas both the physician and nurse cognitive scales are somewhat dependent on styles of documentation in the medical record, the technical diagnostic, technical therapeutic, and intensive care or telemetry monitoring scales are much less dependent on styles of documentation.

To produce these scales, we combined some process measures applicable to all patients with those applicable to subsets of patients. Sicker patients and those with longer hospital stays had a greater number of applicable process criteria than did less-sick patients. In general, compliance with criteria that were applicable only to sicker patients was lower than compliance with criteria that were applicable to all patients. To avoid a bias when combining criteria to form scales, we standardized all process criteria to have a mean of 0 and an SD of 1. The overall process scale represents an average of the five subscales. A patient who underwent an average process of care has an expected process score of 0 and an SD of 1.

To validate our process scales we used logistic regression to examine the relationship between in-hospital process scale scores and mortality 30 and 180 days after admission after adjusting for disease-specific sickness at admission.[14] Linear regression was also used to determine the association of the PPS with change in process.

RESULTS

Reliability and Validity of Measures

Compliance was high for most of the explicitly stated process criteria (Table 1). However, for 21% of our patients with congestive heart failure, 16% of our patients with acute myocardial infarction, and 24% of our patients with pneumonia, the presence or absence of a heart murmur was not noted in the medical record. For 19% of the patients with congestive heart failure, 26% of the patients with pneumonia, and 17% of the patients with cerebrovascular accidents, fewer than three blood pressure readings were taken on day 2 of the hospitalization. Five percent of the patients with congestive heart failure, 6% of the patients with acute myocardial infarction, 9% of the patients with cerebrovascular accidents, 10% of the patients with pneumonia, and 10% of the patients with hip fractures did not have a serum potassium study done on day 1 or 2 of the hospital stay. One fourth of the patients sick enough to be hospitalized for congestive heart failure did not have a serum creatinine study done in the first 2 days, while one third of the patients with congestive heart failure admitted in a moderately sick or very sick condition did not have any creatinine phosphokinase enzyme studies done on day 1 or 2 of the hospitalization to rule out an acute myocardial infarction.

For patients hospitalized with congestive heart failure, acute myocardial infarction, pneumonia, or cerebrovascular accident, better process is significantly associated with a lower 30-day mortality rate. For patients with congestive heart failure, the mortality rate 30 days after admission, adjusted for sickness at admission, was 11% for patients who experienced good process of care, 13% for those who experienced medium process, and 19% for those who

Table 2.—Relationships Between Mortality Rates After Admission Adjusted for Sickness at Admission and Overall Process Scale for Five Diseases

Disease	Mortality Rates 30 Days After Admission, Adjusted for Sickness at Admission,[14] by Overall Process Scale Score Category, %*			P†	Relative Risk of Adjusted 30-Day Death for Poor Compared With Good Care Process‡
	Good	Medium	Poor		
Congestive heart failure	10.7	12.9	18.6	<.01	1.74 (0.23)
Acute myocardial infarction	23.9	22.0	30.1	<.01	1.26 (0.11)
Pneumonia	14.8	15.2	20.2	<.01	1.36 (0.16)
Cerebrovascular accident	18.7	20.3	25.5	<.01	1.36 (0.14)
Hip fracture	5.1	5.2	4.6	>.05	0.90 (0.22)

*Patients were rank-ordered according to process scale scores. Patients with process scale scores in the highest 25% were considered to have experienced good process, those with scores in the lowest 25% poor process, and the remainder medium process.

†From tests of the significance of the process coefficients in the logistic regressions of mortality on process and sickness at admission.

‡Values in parentheses are approximate SEs.

Table 3.—Relationships Between Mortality Rates 30 Days After Admission Adjusted for Sickness at Admission and Process Scales for Five Diseases

Process Subscale	Mortality Rates 30 Days After Admission, Adjusted for Sickness at Admission,[14] by Process Scale Score Category, %*									
	Congestive Heart Failure		Acute Myocardial Infarction		Pneumonia		Cerebrovascular Accident		Hip Fracture	
	Good Process	Poor Process	Good Process	Poor Process	Good Process	Poor Process	Good Process	Poor Process	Good Process	Poor Process
Physician cognitive	12	16†	23	28†	15	19†	18	24†	6	5
Nurse cognitive	11	17†	24	27†	15	19†	19	24†	4	6
Technical diagnostic	11	16†	24	29†	14	19†	19	25†	4	5
Technical therapeutic	11	21†	29	21‡	15	21†	§	§	5	5
Monitoring with intensive care and telemetry	18	13	21	28†	19	15	23	21	10	5‡
Overall	11	19†	24	30†	15	20†	19	26†	5	5

*Patients were rank-ordered according to process scale scores. Patients with process scale scores in the highest 25% were considered to have experienced good process, and those with scores in the lowest 25% poor process.

†$P<.05$ using logistic regression of mortality, adjusted for sickness at admission, on process.

‡Paradoxical $P<.05$—a better process was associated with a worse outcome.

§The technical therapeutic scale was not measured for cerebrovascular accident.

Table 4.—Process Scores Before and After Introduction of the PPS*

Process Subscale	Congestive Heart Failure			Acute Myocardial Infarction			Pneumonia		
	Change in Process Score After PPS†	Expected Change in Mortality Rates After PPS, Percentage Points‡		Change in Process Score After PPS†	Expected Change in Mortality Rates After PPS, Percentage Points‡		Change in Process Score After PPS†	Expected Change in Mortality Rates After PPS, Percentage Points‡	
		30-Day	180-Day		30-Day	180-Day		30-Day	180-Day
Physician cognitive	+0.31§	−0.5	−0.7	+0.24§	−0.7	−0.7	+0.24§	−0.4	−0.5
Nurse cognitive	+0.36§	−0.8	−0.6	+0.22§	−0.6	−0.5	+0.42§	−0.8	−1.3
Technical diagnostic	+0.26§	−0.4	−0.0	+0.21§	−0.6	−0.5	+0.23§	−0.4	−0.5
Technical therapeutic	+0.09§	−0.2	−0.1	+0.16§	+0.2¶	+0.3¶	+0.15§	−0.3	−0.4
Monitoring with intensive care and telemetry	+0.21§	−0.2‖	−0.9‖	+0.05	−0.1	−0.1	+0.08	+0.1‖	−0.0‖
Overall process	+0.42§	−1.2	−1.0	+0.27§	−0.8	−0.7	+0.43§	−1.0	−1.6

*PPS indicates prospective payment system.
†Scores are rated on a scale with a mean of 0 and an SD of 1.
‡Mortality rates are adjusted for sickness at admission.
§$P<.05$ for change in process score after the introduction of the PPS.
‖This expected change in mortality rate is included for completeness; however, the process-outcome link was not sufficiently strong for this process scale to accurately predict a change in the mortality rate from the change in the process score.
¶Paradoxical $P<.05$ for the process-outcome relationship—a better process was associated with a worse outcome.
#The technical therapeutic scale was not measured for cerebrovascular accident.

experienced poor process ($P=.0002$). The relative risk of adjusted 30-day death as process changed from good to poor ranged from 1.74 for congestive heart failure to 1.26 for acute myocardial infarction ($P<.05$, Table 2). We were unable to demonstrate a process-outcome link for patients with hip fractures, partly because 5% of patients with hip fractures died, and this low death rate limited our power to detect a process-outcome relationship.

In addition, a significant process-outcome relationship existed for four of the five process subscales for congestive heart failure, acute myocardial infarction, and pneumonia and for three of the four process subscales for cerebrovascular accidents (Table 3). We found a clinically sensible process-outcome link for the monitoring with intensive care and/or telemetry subscale only for patients with acute myocardial infarction; we defined the need for such monitoring more precisely for acute myocardial infarction than we did for the other diseases.

Process of Care Before and After Introduction of the PPS

For each process scale, for all five diseases, we found better process of care after the introduction of the PPS (Table 4). In all instances the improvement was significant ($P<.05$), except for monitoring with intensive care and/or telemetry, for which the process changed significantly ($P<.05$) only for congestive heart failure and hip fracture. The improvements in process after the introduction of the PPS were apparent both for process measures that could have been influenced by changes in documentation in the medical record (eg, the physician and nurse cognitive scales) and for process measures that were unlikely to be affected by such potential biases (eg, the technical diagnostic and technical therapeutic scales).

We used the previously demonstrated process-outcome link to translate the better process of care after the introduction of the PPS into mortality reductions. For example, for patients with congestive heart failure, the improvement in the process of care of 0.31 SD on the physician cognitive process scale was associated with an expected 0.5 percentage point reduction in the 30-day postadmission mortality rate and an expected 0.7 percentage point reduction in the 180-day postadmission mortality rate. Similar improvements in process on the nurse cognitive scale were associated with expected decreases in mortality of 0.8 and 0.6 percentage points at 30 and 180 days, respectively. Except for hip fracture, the improvements in the overall process scale after the introduction of the PPS were associated with an expected reduction of 0.1 to 1.4 percentage points in the 30-day mortality rate and an expected reduction of 0.4 to 1.6 percentage points in the 180-day mortality rate. Aggregating across our four medical diseases, process improvements after the introduction of the PPS were associated with a 1.0 percentage point reduction in the expected 30-day mortality rate (95% confidence limits, 0.6 to 1.4 percentage points). Given that the observed raw 30-day mortality rate for our four medical diseases was 18.7%, the 1.0 percentage point change represents a 5.3% decline in expected mortality associated with the improvements in process.

The improvements in process scale scores paralleled those found in individual items (Table 1). For example, 58% of patients with congestive heart failure had documentation of a day 2 lung examination before the introduction of the PPS compared with 71% after the introduction of the PPS. Nurses documented at least three blood pressure readings on day 2 for 78% of patients with congestive heart failure before the introduction of the PPS compared with 84% after the introduction of the PPS. The use of oxygen (or intubation) on day 1 for hypoxic patients (ie, pO_2 <60 mm Hg) improved from 87% before the introduction of the PPS to 93% after the introduction of the PPS.

COMMENT

We have demonstrated the validity of our process scales by establishing process-outcome links. If our process scores only reflected recording rather than what happened to patients, we would have been unable to find a statistically significant relationship between better processes of care and lower mortality. In addition, if we were measuring only improvements in recording after the introduction of the PPS vs before the PPS, we would have found improvements in process after the introduction of the PPS only for those process measures that depend heavily on recording (eg, physician and nurse process). However, we have demonstrated a significant process-outcome relationship consistently across diseases and across types (ie, recording-sensitive and -insensitive) of process measures. We found the process of care to be better after the introduction of the PPS.

The lack of a consistent process-outcome relationship for scales based on intensive care and telemetry monitoring was disappointing. We believe the

Cerebrovascular Accident			Hip Fracture		
Change in Process Score After PPS†	Expected Change in Mortality Rates After PPS, Percentage Points‡ 30-Day	180-Day	Change in Process Score After PPS†	Expected Change in Mortality Rates After PPS, Percentage Points‡ 30-Day	180-Day
+0.36§	−0.8	−0.4	+0.16§	+0.1‖	−0.0‖
+0.46§	−0.8	−0.5	+0.31§	−0.1‖	−0.3‖
+0.25§	−0.6	−0.4	+0.22§	−0.0‖	−0.2‖
#	#	#	+0.29§	−0.1‖	−0.2‖
+0.08	−0.2‖	−0.2‖	−0.04§	−0.1¶	−0.2¶
+0.49§	−1.4	−1.0	+0.41§	−0.0‖	−0.4

problem lies in our imperfect measurement of the *if* in the if-then process statements. We need to better understand how to identify the group of patients for whom use of intensive care and telemetry monitoring makes a difference.[15]

It is notable that we found a significant process-outcome relationship for patients with all four of the medical diseases but not for patients with hip fractures. This may be because short-term mortality occurs less often for patients with hip fractures than for those with medical diseases. Alternatively, mortality may not be the best outcome to study for patients with hip fractures. Another possibility is that the medical record does not provide an adequate data source for evaluating surgical, particularly intraoperative, processes. Methods for better evaluating surgical processes of care are needed.

Our consistent findings across process subscales and diseases suggest that the process of care has improved from 1981 to 1986. The implementation of the PPS was not associated with a deterioration in care, even in those areas that were most sensitive to the financial incentives provided by prospective payment to decrease the level of services, such as nursing activities and the use of intensive care units. This is an encouraging finding and indicates that the 24% decrease in length of stay after the introduction of the PPS was not associated with a deterioration in the process of in-hospital care. If anything, improvements in the process of care after the introduction of the PPS should lead to about a 1 percentage point reduction in the 30-day mortality rate. This effect was produced by improvements in both physician and nursing care.

Our other notable result was the large difference in mortality, adjusted for admission sickness, between patients at the top and bottom ends of our overall process scale. For the four medical conditions combined, the adjusted mortality rate went from 17.0% to 23.6% as the process went from the highest to the lowest quartile, an effect more than six times greater than the effect on expected mortality of process improvements after the introduction of the PPS. Further analyses may identify which types of patients and hospitals tend to fall in the lower or upper process quartiles. With this information an ongoing clinical system for improving care can be developed through which professionals can improve the care they give and thereby improve the health of the American people.

This study was supported by cooperative agreement 18-C-98853/9-03 issued by the Health Care Financing Administration of the US Department of Health and Human Services. The entire costs to the RAND Corp for this project ($3.9 million) were funded with federal money.

We gratefully acknowledge the collaborative effort of the five professional review organizations that enabled this work to be completed. Participants from the professional review organizations included medical directors, physician-consultants and -reviewers, project directors, and review coordinators from each of the five study states. In particular, we appreciate the keen clinical insight of the specialist physicians from the professional review organizations with whom we consulted throughout this study. We also acknowledge the many contributions of Harry Savitt, PhD, project officer from the Office of Research and Demonstrations of the Health Care Financing Administration of the US Department of Health and Human Services, whose administrative skills, astute commentary, and continuing support helped us immeasurably. Stanley S. Bentow, MS; Maureen Carney, MS; Patricia A. Damiano; Carole Chew, RRA, MPH; Linda Ferguson; Caren Kamberg, MSPH; Nancy J. Palmer; and Carol Roth, RN, MPH, were key contributors to project development and implementation. Andrea Steiner, MS, gave valuable editorial advice. Our policy advisory board (John C. Beck, MD; Barbara J. Burns, PhD; Monroe T. Gilmour, MD; Paul F. Griner, MD; Charlene Harrington, RN, PhD; T. Reginald Harris, MD; Rosalie Kane, DSW; Shirley Kellie, MD; Judith R. Lave, PhD; Charles E. Lewis, MD; Joseph Martin; Francis D. Moore, Jr, MD; Richard N. Pierson, Jr, MD; and James F. Rodgers, PhD) provided astute advice and guidance. We acknowledge the important contributions of the 297 hospitals whose medical records we reviewed. Without the efforts of these individuals and institutions, this evaluation could not have been successfully completed. Finally, we thank Florence McGinty, without whose secretarial skills this article would not have been produced.

References

1. Donabedian A. Evaluating the quality of medical care. *Milbank Q.* 1966;44:166-203.
2. Brook RH, Appel RA. Quality of care assessment: choosing a method for peer review. *N Engl J Med.* 1973;288:1323-1329.
3. Wyszewianski L. Quality of care: past achievements and future challenges. *Inquiry.* 1988;25:13-22.
4. McAuliffe WE. Studies of process-outcome correlations in medical audits: a critique. *Med Care.* 1978;16:907-930.
5. Rubenstein LV, Mates S, Sidel VW. Quality of care assessment by process and outcome scoring. *Ann Intern Med.* 1977;86:617-625.
6. Greenfield S, Cretin S, Worthman L. Comparison of criteria maps to a criteria list in quality of care assessment for patients with chest pain: the relation of each to outcome. *Med Care.* 1981;19:255-272.
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, 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 Corp; 1988. N-2798-HCFA.
9. 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 Corp; 1988. N-2802-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 Corp; 1988. N-2799-HCFA.
11. 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 Corp; 1988. N-2800-HCFA.
12. Roth CP, Kahn KL, Sherwood MJ, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Pneumonia.* Santa Monica, Calif: RAND Corp; 1988. N-2801-HCFA.
13. Nunnally JC. Assessment of reliability. In: Nunnally JC, ed. *Psychometric Theory.* New York, NY: McGraw-Hill International Book Co; 1978:193-197.
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. Griner PF. Medical intensive care in the teaching hospital: costs versus benefits: the need for an assessment. *Ann Intern Med.* 1973;78:581-585.

Changes in Quality of Care for Five Diseases Measured by Implicit Review, 1981 to 1986

Lisa V. Rubenstein, MD, MSPH; Katherine L. Kahn, MD; Ellen J. Reinisch, MS; Marjorie J. Sherwood, MD; William H. Rogers, PhD; Caren Kamberg, MSPH; David Draper, PhD; Robert H. Brook, MD, ScD

We measured quality of care before and after implementation of the prospective payment system. We developed a structured implicit review form and applied it to a sample of 1366 Medicare patients with congestive heart failure, acute myocardial infarction, pneumonia, cerebrovascular accident, or hip fracture who were hospitalized in 1981-1982 or 1985-1986. Very poor quality of care was associated with increased death rates 30 days after admission (17% with very good care died vs 30% with very poor care). The quality of medical care improved between 1981-1982 and 1985-1986 (from 25% receiving poor or very poor care to 12%), although more patients were judged to have been discharged too soon and in unstable condition (7% vs 4%). Except for discharge planning processes, the quality of hospital care has continued to improve for Medicare patients despite, or because of, the introduction of the prospective payment system with its accompanying professional review organization review.

(*JAMA*. 1990;264:1974-1979)

QUALITY of care can be judged either by implicit or explicit review.[1-6] Explicit review relies on a priori fixed criteria, while implicit review is dependent on the practitioner's opinions. Implicit review of the medical record is the current community gold standard for making final judgments about the quality of care.[7] The purpose of this study was to improve the implicit review method, to determine its reliability and validity, and then to use it to evaluate changes in the quality of medical care for Medicare patients in the United States between 1981 and 1986.

From the Health Program of the RAND Corp, Santa Monica, Calif (Drs Rubenstein, Kahn, Sherwood, Rogers, Draper, and Brook and Mss Reinisch and Kamberg); and the Departments of Medicine (Drs Rubenstein, Kahn, and Brook) and Health Services (Dr Brook), UCLA.

The opinions, conclusions, and proposals in the text are those of the authors alone and do not necessarily represent the views of the RAND Corp or UCLA.

Reprint requests to the RAND Corp, 1700 Main St, PO Box 2138, Santa Monica, CA 90406-2138 (Dr Rubenstein).

PATIENTS AND METHODS

Patient Sample

The implicit review sample was randomly selected from the 14 012 Medicare patients who were included in the study.[8,9] One thousand three hundred sixty-six medical records (10%) were selected to undergo implicit review. Deaths were oversampled; approximately 50% of patients whose records underwent implicit review had died in the hospital. In the analyses reported herein, data have been reweighted to reflect the original 14 012 patients sampled.

Performing Implicit Review

To perform implicit reviews, reviewers were instructed to examine the entire medical record with the exception of nursing notes.[10] Nursing notes were available to the reviewer for use as needed but because of time constraints were not reviewed in their entirety. Reviewers answered 27 questions that covered the process of physician and nursing care; the appropriateness of use of hospital services; patient prognosis; treatability of the patient's condition; preventability of death when it occurred; the quality of the outcome; and an overall assessment of the quality of care provided during the hospitalization. Ratings were based on Likert scales; a five-point scale from very poor to excellent was used for most of the items. We used the same review form for patients with congestive heart failure, myocardial infarction, and pneumonia (Figure). We modified the form slightly for hip fracture and for cerebrovascular accident.

Twenty-five physician - reviewers participated in the study. One reviewer per disease was selected by each of the five state professional review organizations participating in the study, but each reviewer reviewed records from all states. We randomly assigned records to reviewers. No reviewer reviewed patients from more than one of the five disease groups. All reviewers were board certified. Internists reviewed records of congestive heart fail-

Reprinted from *JAMA, The Journal of the American Medical Association*, October 17, 1990, Vol. 264, No. 15, pp. 1974–1979, © 1990, American Medical Association. Reprinted by permission.

DRG/QC STUDY
IMPLICIT REVIEW FORM
(For Congestive Heart Failure, Acute Myocardial Infarction and Pneumonia)

Case ID: _____ Review Date: ____ (Month) ____ (Day) ____ (Year)

1) Please rate the quality of physician and nurse documentation of each of the following patient's prior and chronic disease, functional status, habits, and psychosocial status prior to the current acute illness.

	excellent (1)	good (2)	adequate (3)	poor (4)	very poor (5)	
a) physician documentation of prior and chronic disease	___	___	___	___	___	15/
b) physician documentation of functional status (e.g., ambulation)	___	___	___	___	___	16/
c) physician documentation of habits (e.g., alcohol, smoking, diet)	___	___	___	___	___	17/
d) physician documentation of psychosocial status (e.g., dementia, depression, nursing home residence)	___	___	___	___	___	18/
e) nurse documentation of prior and chronic disease, functional status, habits, and psychosocial status	___	___	___	___	___	19/

f) Check here if the record demonstrates evidence that the physician has ready access to additional records that supplement the current data regarding the patient's prior condition. ☐ 20/

2) Please rate the physician initial assessment of acute medical problems present at admission. Base your answer on the history, physical, and labs.

	excellent (1)	good (2)	adequate (3)	poor (4)	very poor (5)	
a) completeness of initial data gathering	___	___	___	___	___	21/
b) integration of admission information and development of appropriate diagnoses	___	___	___	___	___	22/
c) initial treatment plan and initial orders	___	___	___	___	___	23/

ANSWER QUESTION 6 ONLY IF THE PATIENT WAS DISCHARGED ALIVE

6) Was length of stay appropriate given the patient's status at discharge and disposition plans?

definitely yes	___ (1)	47/
probably yes	___ (2)	
probably no	___ (3)	
definitely no	___ (4)	

6a) If probably or definitely not appropriate, how would you describe length of stay?

too short	___ (1)	48/
too long	___ (2)	

6b) If length of stay was too short or too long, what were the apparent reasons? Check one or more reasons if applicable.

Too short		Too long		
i) Patient too unstable	☐	v) Waiting for nursing home or ECF bed	☐	49-50/
ii) Work-up incomplete	☐	vi) Waiting for home care support service	☐	51-52/
iii) Rehabilitation incomplete	☐	vii) Patient or family refused discharge	☐	53-54/
iv) Patient or family insisted on discharge	☐	viii) Waiting for procedure	☐	55-56/

ANSWER QUESTION 7 ONLY IF THE PATIENT DIED DURING THE HOSPITALIZATION

7) Was the patient's death preventable?

definitely preventable	___ (1)	57/
probably preventable	___ (2)	
probably not preventable	___ (3)	
definitely not preventable	___ (4)	

3) Assume adequate to optimal care and assume the patient recovers from this episode of acute myocardial infarction. What do you believe is this patient's life expectancy?

< 1 month	___ (1)	24/
1-6 months	___ (2)	
> 6 months-1 year	___ (3)	
> 1 year	___ (4)	

4) Assume adequate to optimal care by the physician and hospital. How effective is medical science in treating this patient's acute illness or in preventing worsened health status due to the illness? Consider the severity of the patient's acute illness and the patient's chronic reserve.

very effective	___ (1)	25/
effective	___ (2)	
not so effective	___ (3)	
very ineffective	___ (4)	

5) Considering the entire hospitalization, on average, was use of these services appropriate with respect to the patient's needs? If not appropriate, was it because of underuse?

	definitely yes (1)	probably yes (2)	probably no (3)	definitely no (4)	underuse	
a) monitoring intensity	___	___	___	___	a)	26/
ai) intensive care	- - - - - - - - - - - - - - - - - - - -				ai)	27/
aii) telemetry without intensive care	- - - - - - - - - - - - - - - - - - - -				aii)	28/
b) respiratory therapy delivered	___	___	___	___	b)	29-30/
c) 0_2 and ventilation	___	___	___	___	c)	31-32/
d) arterial blood gases	___	___	___	___	d)	33-34/
e) physical therapy delivered	___	___	___	___	e)	35-36/
f) EKGs	___	___	___	___	f)	37-38/
g) chest x-rays	___	___	___	___	g)	39-40/
h) venous blood tests, urinalyses, sputum analyses	___	___	___	___	h)	41-42/
i) consultations	___	___	___	___	i)	43-44/
j) medications (type and route)	___	___	___	___	j)	45-46/

ANSWER THE REMAINING QUESTIONS FOR ALL PATIENTS

8) How would you characterize the patient's outcome at discharge?

much better than expected	___ (1)	58
better than expected	___ (2)	
as expected	___ (3)	
worse than expected	___ (4)	
much worse than expected	___ (5)	

9) Considering everything you know about this patient, please rate overall quality of care.

extreme, above standard	___ (1)	59/
above standard	___ (2)	
adequate	___ (3)	
below standard	___ (4)	
extreme, below standard	___ (5)	

10) Would you send your mother to these physicians in this hospital?

definitely yes	___ (1)	60/
probably yes	___ (2)	
probably no	___ (3)	
definitely no	___ (4)	

Implicit review form.

ure, cardiologists reviewed acute myocardial infarction records, pulmonologists reviewed pneumonia records, neurologists reviewed cerebrovascular accident records, and orthopedists reviewed hip fracture records.

We trained reviewers in the use of a structured form with its accompanying written guidelines (Figure).[10] Reviews were budgeted at 30 minutes each and training at 12 hours per physician. In training, we specifically avoided trying to change reviewers' opinions about what should have been done in a given clinical situation, but we encouraged reviewers to use a uniform set of rating terms as applied to predefined aspects of care.

Initial training of physician-reviewers was performed during one 3-hour small-group session. Following the session, each physician reviewed two training records and participated in a single half-hour telephone call. Physicians then reviewed three more records on their own at home; these records were discussed during a subsequent 2-hour conference call involving all five physicians and two study investigators.

Reviewers were taught during training to anchor their ratings. For most items, reviewers were asked to use the lowest response category when the care given would have been highly likely to contribute to a bad outcome. For example, for the quality of a physician's history or clinical assessment, raters were instructed to use the lowest response category (very poor) if the rater, when asked to see the patient at midnight, would have to start from scratch in the evaluation. A judgment of "adequate" meant that most essential historical and assessment observations were included, but additional data might be required for optimal diagnosis and treatment. "Excellent" meant that all necessary data were present.

Reviewers were asked to judge urban, rural, teaching, and nonteaching hospitals according to the same standard. Reviewers were asked to rate care as inadequate when that care did not meet a level achievable by most hospitals and not to judge care as inadequate solely for failure to perform extraordinary or controversial procedures. Reviewers were asked to take account of do not resuscitate orders and not to second-guess the medical record regarding the level of aggressiveness of treatment aspired to by the physicians caring for the patient.

To determine interrater reliability, a randomly selected sample of 25% of records was reviewed by two reviewers, and 3% were reviewed by all five reviewers.

Additional Data Sources

We used data from the previously described explicit process reviews of quality of care[6] and mortality 30 days after hospitalization[5] to study whether two methods of process review produce similar results and to examine the predictive validity of implicit process review.[1]

Statistical Methods

We developed process scales and an overall quality-of-care scale. We used factor analysis to determine whether our process groupings made sense psychometrically; final scales were quite similar to our initial, clinically developed scale groupings. Cases reviewed by multiple reviewers were assigned the mean score across reviewers.

The overall quality-of-care scale was based on answers to two questions: (1) "Considering everything you know about the patient, please rate overall quality of care," with five response categories, ranging from extreme, above standard, to extreme, below standard. (2) "Would you send your mother to these physicians in this hospital?" with four response categories, ranging from definitely yes to definitely no. Responses to these questions were added together to form an 8-point overall quality-of-care scale, ranging from 2 (both questions answered with the highest possible rating of 1) to 9 (worst care). The scale was divided into four levels: (1) "very poor care," a score of 7.5 to 9; (2) "poor care," a score of 6.5 to 7.4; (3) "good care," a score of 3.5 to 6.4; and (4) "very good care," a score of 2 to 3.4. Poor care or very poor care means that responses to both of the overall quality questions were in the "below standard" and "probably would not send my mother" range.

Reliability.—We hypothesized that some reviewers might judge care more harshly and some more leniently. We tested for the significance of this "reviewer effect" using a one-way analysis of variance. We then adjusted for the reviewer effect. We used analysis of variance on the adjusted scores to assess interrater reliability.

Components of Variance. — We used components of variance analysis to understand reasons for disagreement among multiple reviewers of the same case and to examine the relationship between this disagreement and the conclusions that could be drawn about the true quality differences between patients who were assigned particular quality ratings.[11] The figures generated by this method provide a conservative estimate of the amount of care judged to be inadequate.

Evaluating the Relationship Between Quality of Care and Death Within 30 Days.—We used linear regression to evaluate whether poor quality of care measured implicitly by the overall quality scale predicted death within 30 days after adjusting for patient sickness.[12] Logistic regression produced similar results.

RESULTS

Patient Sample

A total of 278 records were reviewed for congestive heart failure, 275 for acute myocardial infarction, 273 for pneumonia, 270 for cerebrovascular accident, and 270 for hip fracture. This represented 93% of records selected.[13]

Table 1.—Comparison of Demographic Characteristics in Implicit and Explicit Review Samples

	Implicit Review (n=1366), %	Explicit Review (n=14 012), %
Patients		
Age ≥80 y	43	41
Female	54	57
Nonwhite	16	19
Hospitals		
Rural	20	21
Any teaching	36	35
County	13	14
Serves high percentage of Medicaid patients	16	19

Of the 1366 records reviewed, 993 were reviewed once, 333 were reviewed twice, 33 were reviewed five times, and seven hip fracture records were reviewed four times. The implicit and explicit samples from which the implicit records were drawn did not differ significantly (Table 1).

Reliability

Process scales and the overall quality scale, with sample items and reliabilities, are listed in Table 2. Cronbach's α or interitem reliabilities were between 0.8 and 0.9. Interrater reliabilities for scales were mostly between 0.4 and 0.7.

Overall Quality-of-Care Scale

The Pearson correlation coefficient between the two questions on the overall quality-of-care scale (Table 2) was 0.86. The overall quality scale was also closely correlated with the four process scales and with the preventable death item. For example, the high correlation between the physician's initial assessment of acute disease and the overall quality scale ($r=.80$) indicated that reviewers were much more positive in their assessments of the overall quality of care when the reviewers perceived that the patient was properly assessed by the physician early in the hospitalization. The correlation of the overall quality-of-care scale with nurse assessment, however, was low ($r=.23$), indicating that our overall quality score did not strongly reflect the reviewer's judgment of the initial nursing assessment. This was as expected, because we did not require review of all nursing notes. The correlation with patient prognosis was also low ($r=.01$), indicating that physicians were not confusing quality-of-care ratings with patient sickness at admission. Responses to all of the process items and scales other than the overall quality scale accounted for 72% of the variance in the overall quality scale.

To validate further our structured implicit review, we examined the rela-

Table 2.—Reliability of Implicit Review Scales and Items for Five Diseases

Implicit Review Scales	Range of No. of Items in Scale Across Five Diseases	Cronbach's Interitem Reliability, Mean (Range*)	Interrater Reliability, Mean (Range*)	Example Questions From Structured Implicit Review Form
Physician's initial assessment of function, habit, and chronic disease	2-4	.88 (.82-.92)	.69 (.65-.74)	Rate physician documentation of prior and chronic disease from "excellent" to "very poor"
Physician's initial assessment of acute disease	2-4	.88 (.67-.94)	.59 (.25-.69)	Rate completeness of initial physician data gathering from "excellent" to "very poor"
Use of laboratory tests during hospitalization	3-4	.80 (.74-.88)	.53 (.22-.74)	Rate use of venous blood tests, urinalyses, and sputum analyses from "definitely appropriate" to "definitely inappropriate"
Initial treatment plan, orders, and use of medical therapies	3-6	.82 (.69-.88)	.49 (.36-.65)	Rate initial physician's treatment plan and orders from "excellent" to "very poor"
Overall quality of care during hospitalization	2	.92 (.87-.94)	.54 (.42-.66)	Considering everything you know about this patient, please rate overall quality of care, from "extreme, above standard" to "extreme, below standard" and from "definitely would send my mother to these physicians in this hospital" to "definitely would not send my mother to these physicians"

*The range is given across the five diseases.

tionship between the process of care and mortality and found (Table 3) that, at each level (quartile) of sickness at admission, death within 30 days of hospital admission increased as the quality of care measured by implicit review decreased. For patients in the third quartile of sickness, 13% of people who experienced good or very good care died, while 34% who experienced a very poor care process died. Aggregating across the admission sickness quartiles, the relative risk of death within 30 days for patients with very poor care compared with all other patients was 2.08 ($P<.01$). Reviewers did not appear to lower their implicit judgment of the quality of care for patients who died in the hospital compared with patients who died outside the hospital. After controlling for day of death, the ratings of quality of care for patients who died in and out of the hospital were equivalent and showed similar trends in relation to the 30-day mortality rate.

Table 3.—Patients Who Died Within 30 Days of Hospitalization

Quality of Care by Implicit Review	No. of Patients	Patients Who Died by Level of Sickness at Admission, %*				
		1 (Least Sick) (n=294)	2 (n=301)	3 (n=300)	4 (Most Sick) (n=302)	All Patients (n=1197)
Very good	189	1	5	13	46	17
Good	675	3	6	13	34	14
Poor	164	2	4	8	49	13
Very poor	169	6†	17‡	34‡	55‡	30‡
Total	**1197**	3	7	15	41	17
Relative risk of death within 30 days for patients with very poor quality of care compared with all other patients	...	2.43†	3.09‡	2.79‡	1.42‡	2.08‡

*The level of sickness at admission was measured explicitly.[12] We constructed severity quartiles for each of the five diseases by dividing patients into four equal groups, ranging from the least sick 25% to the most sick 25%.
†$P<.05$.
‡$P<.01$.

Comparison of Implicit and Explicit Reviews

Explicit process scale scores were consistently lower (worse) for patients who received poor or very poor care as judged by implicit review (Table 4). We provide as examples the results comparing the implicit overall quality of care ratings with three of the explicit process scales. These three explicit process scales measured whether a physician recorded specified information about the history and physical examination on day 1 or 2 of the hospitalization, whether a physician recorded such information on day 3, and whether highly abnormal laboratory results were ever rechecked during the hospitalization. There were 67 possible comparisons between explicit and implicit scores. In 54 cases (81%), patients rated good or very good on implicit review had higher explicit scores than patients rated poor or very poor on implicit review; in nine cases (13%), patients had equivalent scores; and in the remaining four cases (6%), patients rated good or very good on implicit review had lower explicit scores than patients rated poor or very poor on implicit review ($P<.001$).

Overall Quality of Care and Differences After the PPS Was Introduced

Reviewers judged the quality of care for our study diseases to be generally good; 82% of patients (all five diseases and all years combined) were considered to have received good or very good care and 18% to have received poor or very poor care.

For all diseases, reviewers judged the quality of care during 1985-1986 more favorably than during 1981-1982 (Table 5). This difference was statistically significant for acute myocardial infarction and cerebrovascular accident ($P<.001$) and for congestive heart failure ($P<.01$). The trend was not significant for pneumonia ($P<.1$) or for hip fracture ($P<.2$). Virtually all scales and items on the form were rated higher in 1985-1986, after the prospective payment system (PPS) was introduced, and most differences between the time periods were significant at $P<.001$ when we aggregated our findings across the five diseases. The number of patients in unstable condition at discharge, however, was assessed by reviewers as significantly worse during 1985-1986 than during 1981-1982 (Table 6).

Physicians judged length of stay to be more appropriate overall in 1985-1986 than in 1981-1982 for three diseases (acute myocardial infarction, pneumo-

nia, and cerebrovascular accident) and about equally appropriate for two diseases (congestive heart failure and hip fracture). The distribution of stays that were inappropriately too long and inappropriately too short changed after the PPS was introduced, however, with fewer inappropriate stays during 1985-1986 due to inappropriately long stays and more due to inappropriately short stays. This was particularly true of congestive heart failure, for which 32% of patients were judged to have been discharged too soon during the later period compared with 24% in the earlier period, and of hip fracture, for which 23% of patients were judged to have been discharged too soon in the later period compared with 12% in the earlier period. For these two diseases, in 1985-1986, after the PPS was introduced, nearly half of the patients who were discharged too soon were judged to have been in unstable condition. Overall, the fraction of hospitalizations rated both as too short and as resulting in patients discharged in unstable condition almost doubled from after the PPS was introduced (4.0% vs 7.1%, $P<.05$).

Table 4.—Comparison Between Implicit and Explicit Process Scale Scores for Five Diseases

Overall Implicit Quality Scale	Average Explicit Process Scale Scores*: Was Specified Information About the History and Physical Examination Recorded on Days 1 or 2? (n=1849)†	Was Specified Information About the History and Physical Examination Recorded on Day 3? (n=1316)‡§	Were Abnormal Laboratory Results Ever Rechecked? (n=909)‖
Very good care	88	76	73
Good care	75	59	68
Poor care	69	48	57
Very poor care	63	47	55

*A higher score (0 to 100) indicates better performance.
†$P<.001$ for all pairwise comparisons.
‡This question was not addressed for hip fractures.
§$P<.001$ for all pairwise comparisons except for poor vs very poor, which was not significant.
‖$P<.01$ for very good vs poor or very poor and for good vs very poor.

Table 5.—Patients With Poor or Very Poor Care Before (1981-1982) and After (1985-1986) the Introduction of the PPS*

Disease	No. of Patients	Patients Who Received Poor or Very Poor Care, %†: Before PPS (1981-1982) (n=669)	After PPS (1985-1986) (n=697)	Difference Between 1981-1982 and 1985-1986‡
Congestive heart failure	278	35	22	−13§
Acute myocardial infarction	275	33	13	−20‖
Pneumonia	273	17	10	−7¶
Cerebrovascular accident	270	36	16	−20‖
Hip fracture	270	5	2	−3
All five diseases	1366	25	12	−13‖

*PPS indicates prospective payment system.
†Reviewer effects were removed (see the "Methods" section).
‡P values take into account not only the size of the differences between 1981-1982 (before the PPS) and 1985-1986 (after the PPS) but also the interitem reliability of the overall quality scale.
§$P<.01$.
‖$P<.001$.
¶$P<.1$.

COMMENT

Over the past 20 years, implicit review of the medical record by physicians has been the community gold standard for judging medical care.[7] Research has demonstrated that this approach to assessment of the quality of care has value,[14-20] but few studies have rigorously evaluated the implicit review technique.[2,14,15] We retooled this classic method for performing quality review and found that our structured approach to implicit review is as reliable as many other clinical measurements physicians perform, such as determining the presence or absence of the dorsalis pedis pulse (39% agreement) or determining whether an electrocardiogram is abnormal (84% agreement).[21,22] We also demonstrated a process-outcome link with implicit review (eg, higher scores on implicit review lead to better outcomes). We used a nationally representative and therefore diverse sample of physicians and patients.[9,23] Reliabilities were sufficiently high to allow for meaningful comparisons between groups of patients, though a single review may not be reliable enough to judge accurately

Table 6.—Patients Judged to Have Inappropriate Length of Stay for Five Diseases Before (1981-1982) and After (1985-1986) the Introduction of the PPS*

Disease	No. of Patients	Patients Who Probably or Definitely Had an Inappropriate Length of Stay, %†: Too Long or too Short a Stay: Before PPS	Too Long or too Short a Stay: After PPS	Too Long a Stay: Before PPS	Too Long a Stay: After PPS	Too Short a Stay: Before PPS	Too Short a Stay: After PPS	Too Short a Stay and in Unstable Condition at Discharge: Before PPS	Too Short a Stay and in Unstable Condition at Discharge: After PPS
Congestive heart failure	247	34	35	10	3‡	24	32	12.5	17.3
Acute myocardial infarction	212	30	11‡	16	3‡	14	8	2.1	0.0
Pneumonia	235	16	12	10	5‡	6	7	0.6	5.1‡
Cerebrovascular accident	216	42	25‡	32	17‡	10	8	0.6	1.6
Hip fracture	258	27	28	15	5‡	12	23‡	3.6	9.4
All five diseases	1168	30	23‡	17	7‡	13	17	4.0	7.1‡

*PPS indicates prospective payment system.
†Includes only patients discharged alive.
‡$P<.05$.

whether an individual patient received poor care.

We compared implicit and explicit reviews to determine whether these two methods measured the same quality-of-care constructs or whether they measured entirely different aspects of the quality of medical care. Happily, the two methods appear to measure similar concepts.

When we used implicit review to judge the quality of medical care provided to a nationally representative sample of patients with five diseases, we found that, although most care is judged to be good or very good, a significant proportion of care was judged to be poor or very poor. Associated with the institution of prospective payment and implicit review of medical records by professional review organizations, the proportion of care rated poor or very poor decreased between 1981-1982 and 1985-1986 from about one fourth of patients to about one eighth of patients. A similar result was reported by Kahn et al[6] using explicit review.

Many fewer patients were judged to have been kept in the hospital too long during the later period. Of concern, however, is the increasing number of patients judged to have been discharged too soon and, in particular, in unstable condition. A similar increase in patients discharged in unstable condition was reported by Kosecoff et al[24] using explicit review criteria.

In summary, structured implicit review performed well as a method for measuring the quality of medical care provided to Medicare patients. This is good news for professional review organizations, which perform the bulk of quality reviews nationally and use implicit review to do so. The study's optimistic result that the quality of medical care improved after the introduction of the PPS should serve as encouraging news—hospitals, physicians, and other health professionals continue to put patient needs first. However, substantial improvement is still possible, and the increasing rate of discharges with patients in unstable condition found by both implicit and explicit review is of concern.

This work was supported by cooperative agreement 18-C-98853/9-03 issued by the Health Care Financing Administration of the US Department of Health and Human Services, Baltimore, Md.

We gratefully acknowledge the collaborative effort of the five professional review organizations that enabled this work to be completed. Participants from the professional review organizations included medical directors, physician-consultants and -reviewers, project directors, and review coordinators from each of the five study states. In particular, we appreciate the keen clinical insight of the specialist physicians from the professional review organizations with whom we consulted throughout this study. We also acknowledge the many contributions of Harry Savitt, PhD, project officer from the Office of Research and Demonstrations of the Health Care Financing Administration of the US Department of Health and Human Services, whose administrative skills, astute commentary, and continuing support helped us immeasurably. Harris Allen, PhD, and Maureen Carney, MS, were key contributors to project development and implementation. Andrea Steiner, MS, gave valuable editorial advice. Our policy advisory board (John C. Beck, MD; Barbara J. Burns, PhD; Monroe T. Gilmour, MD; Paul F. Griner, MD; Charlene Harrington, RN, PhD; T. Reginald Harris, MD; Rosalie Kane, DSW; Shirley Kellie, MD; Judith R. Lave, PhD; Charles E. Lewis, MD; Joseph Martin; Francis D. Moore, Jr, MD; Richard N. Pierson, Jr, MD; and James F. Rodgers, PhD) provided astute advice and guidance. We acknowledge the important contributions of the 297 hospitals whose medical records we reviewed. Without the efforts of these individuals and institutions, this evaluation could not have been successfully completed. Finally, we thank Florence McGinty, without whose secretarial skills this article would not have been produced.

References

1. Brook RH. *Quality of Care Assessment: A Comparison of Five Methods of Peer Review.* Washington, DC: US Dept of Health, Education, and Welfare; 1973. Publication HRA-74-3100.
2. Altman I, Anderson AJ, Barker K. *Methodology in Evaluating the Quality of Medical Care.* Pittsburgh, Pa: University of Pittsburgh Press; 1969.
3. Donabedian A. *The Definition of Quality and Approaches to its Assessment: Explorations in Quality Assessment and Monitoring.* Ann Arbor, Mich: Health Administration Press; 1980.
4. Donabedian A. *Explorations in Quality Assessment and Monitoring: The Criteria and Standards of Quality.* Ann Arbor, Mich: Health Administration Press; 1982.
5. 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;264:1983-1987
6. 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; 264:1968-1972.
7. Dans PE, Weiner JP, Otter SE. Peer review organizations: promises and potential pitfalls. *N Engl J Med.* 1985;313:1131-1137.
8. Kahn KL, Brook RH, Draper D, et al. Interpreting hospital mortality data: how can we proceed? *JAMA.* 1988;260:3625-3628.
9. 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.
10. Kahn KL, Rubenstein LV, Sherwood MJ, Brook RH. *Structured Implicit Review for Physician Measurement of Quality of Care: Development of the Form and Guidelines for Its Use.* Santa Monica, Calif: RAND Corp; 1989.
11. Nunnally JC. *Psychometric Theory.* New York, NY: McGraw-Hill International Book Co; 1978.
12. 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.
13. Kahn KL, Draper D, Keeler EB, et al. *The Effects of DRG-Based Prospective Payment on Quality of Care for Hospitalized Medicare Patients.* Santa Monica, Calif: RAND Corp; 1990.
14. Dubois RW, Moxley JH, Draper D, Brook RH. Hospital inpatient mortality: is it a predictor of quality? *N Engl J Med.* 1987;317:1674-1680.
15. Richardson FM. Peer review of medical care. *Med Care.* 1972;10:29-39.
16. Morehead MA, Donaldson RS, Seravalli MR. Comparisons between OEO neighborhood health centers and other health care providers of ratings of the quality of health care. *Am J Public Health.* 1971;61:1294-1305.
17. Payne BC. Continued evolution of a system of medical care appraisal. *JAMA.* 1967;201:536-540.
18. Peters EN. Practical versus impractical peer review. *Med Care,* 1972;10:516-521.
19. Sheps MC. Approaches to the quality of hospital care. *Public Health Rep.* 1955;70:877-886.
20. Fine J, Morehead MA. Study of peer review of inhospital patient care. *NY State J Med.* 1971;71:1963-1973.
21. Koran LM. The reliability of clinical methods, data and judgments, I. *N Engl J Med.* 1975; 293:642-646.
22. Koran LM. The reliability of clinical methods, data and judgments, II. *N Engl J Med.* 1975;293:695-701.
23. Wennberg J, Gittelsohn A. Small-area variations in health care delivery. *Science.* 1973; 182:1102-1108.
24. Kosecoff J, Kahn KL, Rogers WH, et al. Prospective payment system and impairment at discharge: the 'quicker-and-sicker' story revisited. *JAMA.* 1990;264:1980-1983.

Prospective Payment System and Impairment at Discharge

The 'Quicker-and-Sicker' Story Revisited

Jacqueline Kosecoff, PhD; Katherine L. Kahn, MD; William H. Rogers, PhD; Ellen J. Reinisch, MS; Marjorie J. Sherwood, MD; Lisa V. Rubenstein, MD, MSPH; David Draper, PhD; Carol P. Roth, RN, MPH; Carole Chew, RRA, MPH; Robert H. Brook, MD, ScD

Since the introduction of the prospective payment system (PPS), anecdotal evidence has accumulated that patients are leaving the hospital "quicker and sicker." We developed valid measures of discharge impairment and measured these levels in a nationally representative sample of patients with one of five conditions prior to and following the PPS implementation. Instability at discharge (important clinical problems usually first occurring prior to discharge) predicted the likelihood of postdischarge deaths. At 90 days postdischarge, 16% of patients discharged unstable were dead vs 10% of patients discharged stable. After the PPS introduction, instability increased primarily among patients discharged home. Prior to the PPS, 10% of patients discharged home were unstable; after the PPS was implemented, 15% were discharged unstable, a 43% relative change. Efforts to monitor the effect of this increase in discharge instability on health should be implemented.

(*JAMA*. 1990;264:1980-1983)

SINCE the introduction in 1983 of the prospective payment system (PPS), there has been considerable anecdotal evidence that patients are leaving the hospital "quicker and sicker."[1-4] However, there has been a dearth of systematic evidence to support the assertion that, post-PPS, patients are discharged with an increased level of illness.

We do know that changing the way a hospital was paid was initially associated with dramatic decreases in length of stay,[5] although these decreases have now stabilized.[6] The purpose of this article is to provide nationally representative data about the level of impairment at discharge prior to and following the implementation of the PPS. In order to answer correctly the question of whether patients are inappropriately discharged with an increased level of illness since the implementation of the PPS, we must first define a valid measure of impairment at discharge and then determine if patients discharged from the hospital with an increased level of illness do worse than expected. If patients who are discharged with clinical impairments die more frequently than expected, then increasing the level of illness at discharge might be unwise.

From Value Health Sciences Inc, Santa Monica, Calif (Dr Kosecoff and Mss Roth and Chew); the Health Program of the RAND Corp, Santa Monica, Calif (Drs Kahn, Rogers, Sherwood, Rubenstein, Draper, and Brook, and Ms Reinisch); and the Departments of Medicine (Drs Kosecoff, Kahn, Rubenstein, and Brook) and Health Services (Drs Kosecoff and Brook), UCLA.

The opinions, conclusions, and proposals in this article are those of the authors alone and do not necessarily represent the views of any of these organizations, the RAND Corp, or UCLA.

Reprint requests to Value Health Sciences Inc, 1448 15th St, #202, Santa Monica, CA 90404 (Dr Kosecoff).

METHODS

The study methods, including design, sampling, and analysis strategy, are described in another article in this series.[7] This report, because it concerns patients discharged alive from the hospital, excludes patients who died in the hospital, who were directly transferred to other acute care hospitals, or who had a "do not resuscitate" order written at some time during the hospitalization, reducing the maximum sample size from 14 012 to 10 913.

DISCHARGE IMPAIRMENT: CONSTRUCTION OF MEASURES

Based on clinical judgment, we constructed three measures of discharge impairment (Table 1): instability at discharge, sickness at discharge, and abnormal last laboratory values. Data used to construct the first two measures came from the day of discharge or the day prior to discharge.

Instability variables are designed to identify patient problems present at discharge that (1) clinicians generally agree should be either corrected prior to discharge or monitored in the postdischarge period, and (2) may result in poor outcomes if not corrected. In general, instability variables reflect clinical problems that were not present at admission. Sickness variables are designed to measure sickness at discharge regardless of whether the problem was present at admission or should or could have been corrected at discharge. As seen in Table 1, 8.3% of patients with acute myocardial infarction had sickness at discharge caused by the presence of chest pain on at least one of the last 2 days of hospitalization, but 0.3% of patients with myocardial infarction had instability at discharge caused by chest pain that was not present at admission. Abnormal last laboratory values reflect the presence of abnormal findings the last time a test was performed during the hospitalization.

ANALYSIS

Virtually all of the analyses presented herein are based on χ^2 tests of significance. We did, however, use multivariate techniques (linear and logistic regression) to adjust for differences in patient sickness at time of hospital admission when studying discharge impairment pre- and post-PPS. We also studied the relationship between discharge impairment and postdischarge death separately in the pre- and post-PPS periods and defined death within a specified number of days either postadmission or postdischarge.[8,9] The results from these analyses are qualitatively consistent with the analyses reported herein. In this article, emphasis is given to the instability at discharge measure because it is clinically the most appealing of our three measures.

RESULTS

Considering all study years, one (17%) of six patients was discharged with at least one instability, two (39%) of five patients were discharged with at least one measure of sickness, and one

Table 1.—Definition of Measures of Discharge Impairment

Measures of:	% of Patients Discharged With a Measure of Impairment*					
	CHF (n=2348)	AMI (n=1946)	PNE (n=2141)	CVA (n=1927)	HIP (n=2551)	5 Diseases (n=10 913)
Instability at discharge†						
Fever, temperature >38.3°C	0.0	0.2	0.2	0.3	0.2	0.2
New incontinence	4.6	. . .	5.8	15.6	12.3	9.5
New chest pain	1.4	0.3	1.3	. . .	2.8	1.5
New shortness of breath	1.5	1.1	1.5	1.5	. . .‡	1.4
New confusion	4.5	2.8	2.5	. . .	5.5	3.9
New heart rate ≥130 beats/min	0.7	0.2	0.5	0.4	0.2	0.4
New respiratory rate ≥30/min	1.9	0.7	1.8	1.8	1.5	1.6
Diastolic blood pressure ≥105 mm Hg	1.5	0.5	1.1	2.9	0.6	1.3
New systolic blood pressure <90 mm Hg	2.1	1.6	1.2	1.4	0.7	1.4
New low heart rate <50 beats/min	. . .	1.8	. . .	. . .	. . .	. . .
New premature ventricular contractions	. . .	0.9	. . .	. . .	. . .	. . .
≥1 instability	15.4	9.5	14.3	21.3	21.0	16.5
Mean No. of instabilities	0.18	0.10	0.16	0.24	0.24	0.18
Mean No. of instabilities among patients with ≥1 instability	1.2	1.1	1.1	1.1	1.1	1.1
Sickness at discharge†						
Fever, temperature >38.3°C	0.0	0.2	0.2	0.3	0.2	0.2
New incontinence	4.6	. . .	5.8	15.6	12.3	9.5
Chest pain	5.5	8.3	4.8	. . .	3.3	5.3
Shortness of breath	29.5	7.8	15.8	3.7	. . .‡	15.0
Confusion	13.2	7.6	11.1	. . .	10.4	10.7
Heart rate ≥130 beats/min	1.2	0.4	0.8	0.4	0.2	0.6
Respiratory rate ≥30/min	5.5	1.3	8.5	3.3	2.4	4.2
Diastolic blood pressure ≥105 mm Hg	1.5	0.5	1.1	2.9	0.6	1.3
Systolic blood pressure <90 mm Hg	4.4	3.6	1.9	1.6	0.9	2.4
New decubitus	1.2	. . .	2.1	2.4	8.6	3.8
Low heart rate <50 beats/min	. . .	4.1	. . .	. . .	. . .	. . .
Premature ventricular contraction	. . .	2.3	. . .	. . .	. . .	. . .
Parenteral antibiotics	. . .	. . .	36.1	. . .	. . .	. . .
≥1 sickness	47.6	29.4	60.4	25.2	30.8	39.0
Mean No. of sicknesses	0.67	0.36	0.88	0.30	0.39	0.52
Mean No. of sicknesses among patients with ≥1 sickness	1.4	1.2	1.5	1.2	1.3	1.3
Abnormal last laboratory values§						
Potassium‖	6.3	3.8	7.8	7.1	6.7	6.3
Sodium¶	2.6	2.2	2.9	1.8	2.8	2.5
Renal distress#	4.1	2.4	2.6	1.9	1.6	2.5
Low hematocrit**	2.8	. . .	0.5	2.1	3.9	2.3
High WBC††	9.3	. . .	12.7	11.0	13.8	11.7
Weight increase >1.35 kg	2.9	2.2	. . .	. . .	. . .	2.6
CHF by roentgenographic worsening	4.3	0.9	7.0	. . .	4.2	4.1
≥1 laboratory abnormality	27.6	10.7	29.2	20.9	29.1	24.1
Mean No. of abnormalities	0.32	0.11	0.33	0.24	0.33	0.27
Mean No. of abnormalities among patients with ≥1 abnormality	1.16	1.05	1.14	1.14	1.13	1.13

*CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.
†All measures of discharge instability and sickness were evaluated for the presence of symptoms and/or signs on the day prior to discharge and/or the day of discharge.
‡For hip fracture, chest pain is defined as chest pain or shortness of breath.
§Each reported value is the last in-hospital value. When comparisons are made, the last value is compared with the most recent prior in-hospital value. If only one value is available during the hospital stay, that value is counted as the last laboratory value even if it was also the first.
‖Abnormal potassium (milliequivalents per liter) is defined as a value ≤2.9, a value ≥6.0, or a value between 5.5 and 5.9 but rising or 3.0 to 3.4 and falling.
¶Abnormal sodium (milliequivalents per liter) is defined as a value <119, a value ≥155, or a value between 120 and 129 but falling or between 150 and 154 but rising.
#Renal distress (milligrams per deciliter) is defined as a creatinine value >6.0, a rise in creatinine >0.5, a blood urea nitrogen value >100, or a blood urea nitrogen rise ≥ 25.
**Low hematocrit (percent) is defined as <24.9 or between 25 and 34.9 with a drop ≥7.0 (the drop was not measured for pneumonia).
††High white blood cell count (cells × 1000/mm³) is defined as ≥12 and rising (the rise was not measured for pneumonia).

(24%) of four patients had an abnormal last laboratory value (Table 1). The presence of one or more instabilities at discharge varied by a factor of 2 across diseases. Among those patients with one or more measures of instability, most had just one. The development of urinary incontinence contributed most to labeling a person as unstable at discharge for the four diseases for which it was measured.

The prevalence of sickness at discharge per 100 patients varied by disease from 25 to 48. Incontinence contributed most to the sickness measure for patients with a cerebrovascular accident and hip fracture, while for heart failure and pneumonia it was shortness of breath, and for myocardial infarction it was chest pain. Finally, the presence of one or more abnormal last laboratory values varied from a low of one in nine patients discharged with myocardial infarction to about one in three patients discharged with pneumonia or hip fracture.

Discharge to an institution was asso-

Table 2.—Patients Discharged Home or Institutionalized With One or More Measures of Impairment at Discharge

Measures of Impairment*	Discharge Destination	
	Home	Institution
Instability at discharge, %		
CHF	14.1†	22.4
AMI	9.4	10.8
PNE	12.0†	21.3
CVA	14.2†	32.0
HIP	12.8†	26.3
5 Diseases	12.5†	22.4
Sickness at discharge, %		
CHF	46.2†	57.8
AMI	28.2†	45.3
PNE	58.0†	68.0
CVA	16.8†	37.9
HIP	19.3†	38.3
5 Diseases	33.7†	49.5
Abnormal laboratory, %		
CHF	43.5‡	49.9
AMI	10.0†	18.7
PNE	27.9‡	33.3
CVA	17.7†	26.0
HIP	25.8‡	31.3
5 Diseases	25.0†	31.8
Sample size, No. of patients		
CHF	1966	322
AMI	1736	139
PNE	1596	507
CVA	1121	782
HIP	998	1543
5 Diseases	7417	3293

*CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.

†$P<.01$ for χ^2 test of level of impairment of patients discharged home vs institution.

‡$P<.05$ for χ^2 test of level of impairment of patients discharged home vs institution.

Table 3.—Relationship of Instability at Discharge to Likelihood of Dying Postdischarge*

Time Interval	No Instability	Presence of Instability
0 to 30 d, % dead†		
CHF	6.5‡	10.0
AMI	4.3‡	7.0
PNE	3.3‡	8.1
CVA	4.0‡	8.5
HIP	2.5‡	5.8
5 Diseases	4.0§	7.9
31 to 60 d, % dead		
CHF	4.4§	7.8
AMI	2.8	2.5
PNE	3.1	5.3
CVA	3.1	4.9
HIP	1.8‡	3.7
5 Diseases	3.0‡	4.8
61 to 90 d, % dead		
CHF	4.6	5.1
AMI	2.3	0.6
PNE	2.7	3.2
CVA	2.7‡	6.1
HIP	2.0‡	3.9
5 Diseases	2.9	3.8
0 to 90 d, % dead		
CHF	14.2‡	21.3
AMI	9.1	9.9
PNE	8.9‡	15.8
CVA	9.5	18.2
HIP	6.3‡	12.8
5 Diseases	9.6‡	15.6
Sample size, 0 to 90 d, No. of patients dead		
CHF	1831	334
AMI	1620	171
PNE	1716	285
CVA	1315	329
HIP	1868	485
5 Diseases	8350	1604

*CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.

†Percent of patients dead within the time interval, given that they were alive at the beginning of the interval.

‡$P<.01$ for χ^2 test of differences in death rates among patients with and without instability.

§$P<.05$ for χ^2 test of differences in death rates among patients with and without instability.

ciated with a significantly higher likelihood of discharge impairment, according to all three of our measures (Table 2). On average, 22% of patients discharged to an institution had one or more instabilities, while 13% of those going home were so classified (relative risk of 1.8). The relative risk was highest for hip fracture (2.0) and lowest for myocardial infarction (1.1). Fifty percent of patients discharged to an institution had one or more measures of sickness, while 34% of those discharged home were so affected (relative risk of 1.5). For abnormal last laboratory values, 32% were discharged to an institution with an abnormality compared with 25% discharged home (relative risk of 1.3).

For the five diseases studied, patients discharged with at least one instability had a higher probability of dying postdischarge than patients discharged without any instability (Table 3). Aggregating across diseases, the risk of death at 90 days following discharge was 16% for patients discharged unstable and 10% if instability was not present at discharge (relative risk of 1.6). The highest condition-specific relative risk of death by 90 days postdischarge was for patients with hip fractures (2.0) and the lowest was for patients with myocardial infarction (1.1).

Prior to implementation of the PPS, 15% of patients were discharged unstable (Table 4), while afterward, 18% were discharged unstable, a 22% increase (95% confidence interval for difference; 1.9 to 4.7 percentage points). Most of the increase in instability was concentrated in those people who were discharged home. Institutions did not receive a significantly greater number of unstable patients (22% pre-PPS and 23% post-PPS). Patients who were discharged home post-PPS, however, were 43% more likely to be unstable than prior to the PPS. The largest change occurred among patients with hip fractures (9% of those discharged home were unstable prior to the PPS vs 17% who were discharged after the PPS—a 93% increase).

The increase in instability post-PPS was more notable for vital signs (eg, a heart rate of ≥130 beats per minute at discharge was 0.1% pre-PPS and 0.6% post-PPS, $P<.05$) than for symptoms (eg, new incontinence at discharge was 8.5% pre-PPS and 10.4% post-PPS, $P<.05$). This suggests that the increase was not driven only by changes in styles of recording in the medical record.

COMMENT

Other articles in this series demonstrate that the introduction of the PPS was not associated with a rise in either short-term (30-day) mortality or 6-month mortality. In light of this, how important is the finding that instability at discharge has increased since the advent of the PPS? The answer depends on one's vantage point.

From the patient's and family's point of view, it must be comforting to know that the overwhelming majority of patients (85% post-PPS) discharged home leave the hospital in a stable condition. However, the percentage of patients discharged home with an instability has increased. Because the major causes of these instabilities are clinically important, the amount of support needed at home has increased. Being discharged home in an unstable condition, with confusion or incontinence, is, at best, an inconvenience for the family and it may, at worst, expose the patient to an increased risk of death.

Geriatricians might notice that elderly patients who are being discharged in an unstable condition suffer from those conditions that physicians have been emphasizing for years, namely incontinence and confusion. For example, taking pre- and post-PPS together, 10% of patients were discharged with new incontinence that was not present at admission and 4% had new confusion. These findings raise questions such as "Are increases in instability caused by inappropriately early discharges, too many tests in a shortened hospital stay, incorrect use of new medications, or by changes in nursing practices (eg, fewer nurses per patient and less time available to talk with the patient or monitor incontinence or disorientation)?" The answers to these questions require new data, perhaps clinical trials, and the development of valid disease-specific discharge guidelines.

From a hospital or nursing home perspective, our results are both encouraging and potentially disturbing. Instability among patients who were not assigned do not resuscitate status and who were discharged to nursing homes has not increased post-PPS. On the other hand, the observation that one quarter of nursing home patients are admitted with an instability is a disquieting

Table 4.—Relationship of Prospective Payment System to Instability at Discharge by Discharge Destination*†

	Discharge Destination		
	All Patients	Home	Institution
Unstable at discharge, %			
Pre-PPS			
CHF	13.6	11.7	23.9
AMI	8.5	8.2	11.9
PNE	12.6	10.6	18.6
CVA	19.0	12.2	30.1
HIP	18.8	8.8	25.7
5 Diseases	15.0	10.3	22.0
Post-PPS			
CHF	17.1§	16.3‡	21.2§
AMI	10.6	10.8	8.3
PNE	16.0§	13.4	23.4
CVA	23.4§	16.0	33.5
HIP	23.1‡	17.0‡	26.9
5 Diseases	18.3‡	14.7‡	22.7
Absolute change, %			
Post-PPS – pre-PPS			
CHF	+3.5	+4.6	−2.7
AMI	+2.1	+2.6	−3.6
PNE	+3.4	+2.8	+4.8
CVA	+4.4	+3.8	+3.4
HIP	+4.3	+8.2	+1.2
5 Diseases	+3.3	+4.4	+0.7
Relative change, %			
Post-pre			
CHF	+25	+39	−11
AMI	+25	+32	−30
PNE	+27	+27	+26
CVA	+23	+31	+11
HIP	+23	+93	+4
5 Diseases	+22	+43	+3

*Patients with in-hospital death, do not resuscitate orders, or transfer to acute care hospitals are excluded from this and all other analyses in this article.

†CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.

‡$P<.01$ for χ^2 test of post-pre differences in instability at discharge for all patients, for those discharged home, and for those discharged to an institution.

§$P<0.5$ for χ^2 test of post-pre differences in instability at discharge for all patients, for those discharged home, and for those discharged to an institution.

finding. Is the hospital discharge planning process adequate to handle a smooth transition of these patients to institutions, and is the nursing home reimbursement level sufficient to provide adequate care for them? In particular, do nursing home physicians know of these problems, are they noted in the clinical record sent to the nursing home, and are nursing home nurses and physicians developing adequate strategies to address them?

Our data might justifiably be greeted as an additional burden placed at the hospital administrator's doorstep. That may be the case, but because instability is linked to subsequent death and results in decreased quality of life, this burden needs to be addressed. Learning at a hospital level what causes instability at discharge and how to reduce it should become a major joint activity of the hospital administration and medical staff. Either decreasing instability at discharge to zero or developing postdischarge plans to eliminate the effects of instability at discharge could become part of a hospital's continuous improvement effort. It might also receive emphasis in an accreditation program such as that sponsored by the Joint Commission for Accreditation of Health Care Organizations.

Finally, from a policy perspective, these results raise concerns about what information is needed to monitor a change in health policy and how often data on instability at discharge should be collected. These data are now 4 years out of date, and during the last 4 years, hospitals have received less generous levels of reimbursement. We have presented a problem that has been apparent since at least 1981; patients discharged home in an unstable condition have a greater than average expected risk of dying and probably a need for more family support. In addition, we have demonstrated that the risk of being discharged in an unstable condition has increased since the PPS was implemented. Reducing the effect of instability may be possible even without increasing hospital length of stay, but this will require examining, with the aid of clinical data, the impact of various policy options on the occurrence and outcome of instability.

This study was supported by cooperative agreement 18-C-98853/9-03 issued by the Health Care Financing Administration of the US Department of Health and Human Services, Baltimore, Md. The entire RAND costs ($3.9 million) were funded with federal money.

We gratefully acknowledge the collaborative effort of the five professional review organizations that enabled this work to be completed. Participants from these organizations included medical directors, physician consultants and reviewers, project directors, and review coordinators from each of the five study states. In particular, we appreciate the keen clinical insight of the professional review organizations' physician specialists with whom we consulted throughout this study. We also acknowledge the many contributions of Harry Savitt, PhD, project officer from the Office of Research and Demonstrations of the Health Care Financing Administration of the US Department of Health and Human Services, whose administrative skills, astute commentary, and continuing support helped us immeasurably. Maureen Carney, MS, and Caren Kamberg, MSPH, were key contributors to project development and implementation. Andrea Steiner, MS, gave valuable editorial advice. Our policy advisory board (John C. Beck, MD, Barbara J. Burns, PhD, Monroe T. Gilmour, MD, Paul F. Griner, MD, Charlene Harrington, RN, PhD, T. Reginald Harris, MD, Rosalie Kane, DSW, Shirley Kellie, MD, Judith R. Lave, PhD, Charles E. Lewis, MD, Joseph Martin, Francis D. Moore, Jr, MD, Richard N. Pierson, Jr, MD, and James F. Rodgers, PhD) provided astute advice and guidance. We acknowledge the important contributions of the 297 hospitals whose medical records we reviewed. Without the efforts of these individuals and institutions, this evaluation could not have been successfully completed.

References

1. Coulton CJ. Prospective payment requires increased attention to quality of post-hospital care. *Soc Work Health Care.* 1988;13:19-31.
2. Goodwin ME. Medicare's prospective payment system for hospital: the impact on nursing home quality. *QRB.* 1986;198-201.
3. Coe M, Wilkinson A, Patterson P. *Final Report on the Dependency at Discharge Study.* Beverton, Ore: Northwest Oregon Health Systems; 1986.
4. Neu CR, Harrison SC. *Post Hospital Care Before and After the Medicare Prospective Payment System.* Santa Monica, Calif: RAND Corp; 1988.
5. 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.
6. Witsberger C, Kominsky G. *Recent Trends in Length of Stay for Medicare Surgical Patients.* Santa Monica, Calif: RAND Corp; 1990.
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, Draper D, Keeler EB, et al. *The Effects of DRG-Based Prospective Payment on Quality of Care for Hospitalized Medicare Patients.* Santa Monica, Calif: RAND Corp; 1990.
9. 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.

Comparing Outcomes of Care Before and After Implementation of the DRG-Based Prospective Payment System

Katherine L. Kahn, MD; Emmett B. Keeler, PhD; Marjorie J. Sherwood, MD; William H. Rogers, PhD; David Draper, PhD; Stanley S. Bentow, MS; Ellen J. Reinisch, MS; Lisa V. Rubenstein, MD, MSPH; Jacqueline Kosecoff, PhD; Robert H. Brook, MD, ScD

We compared patient outcomes before and after the introduction of the diagnosis related groups (DRG)-based prospective payment system (PPS) in a nationally representative sample of 14 012 Medicare patients hospitalized in 1981 through 1982 and 1985 through 1986 with one of five diseases. For the five diseases combined, length of stay dropped 24% and in-hospital mortality declined from 16.1% to 12.6% after the PPS was introduced ($P<.05$). Thirty-day mortality adjusted for sickness at admission was 1.1% lower than before (16.5% pre-PPS, 15.4% post-PPS; $P<.05$), and 180-day adjusted mortality was essentially unchanged at 29.6% pre- vs 29.0% post-PPS ($P<.05$). For patients admitted to the hospital from home, 4% more patients were not discharged home post-PPS than pre-PPS ($P<.05$), and an additional 1% of patients had prolonged nursing home stays ($P<.05$). The introduction of the PPS was not associated with a worsening of outcome for hospitalized Medicare patients. However, because our post-PPS data are from 1985 and 1986, we recommend that clinical monitoring be maintained to ensure that changes in prospective payment do not negatively affect patient outcome.

(*JAMA*. 1990;264:1984-1988)

TO EVALUATE whether patient outcomes have changed after the implementation of the diagnosis related groups (DRG)–based prospective payment system (PPS) and the professional review organization system, we conducted a study in which we compared outcomes before and after the PPS was introduced. In this article, we report on in-hospital mortality, mortality 30 and 180 days after admission, discharge to and prolonged stay in a nursing home, and readmission to hospitals.

From the Health Program of the RAND Corp, Santa Monica, Calif (Drs Kahn, Keeler, Sherwood, Rogers, Draper, Rubenstein, and Brook, Mr Bentow, and Ms Reinisch); the Departments of Medicine (Drs Kahn, Rubenstein, Kosecoff, and Brook) and Health Services (Drs Kosecoff and Brook), UCLA; and Value Health Sciences Inc, Santa Monica, Calif (Dr Kosecoff).

The opinions, conclusions, and proposals in this article are those of the authors alone and do not necessarily represent the views of any of these organizations, the RAND Corp, or UCLA.

Reprint requests to the RAND Corp, 1700 Main St, PO Box 2138, Santa Monica, CA 90406-2138 (Dr Kahn).

METHODS

We present the study sample, design, and inclusion criteria elsewhere in this series.[1,2]

We used the medical record as our source of in-hospital mortality information and Health Care Financing Administration files to determine mortality status subsequent to 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 92% of the patients in our sample to the Health Care Financing Administration health insurance master file.

We assessed short-term mortality by studying both in-hospital mortality and death within 30 days of the acute care admission. We chose death within 180 days postadmission as our indicator of medium-term mortality. We used the medical record as the source of both the patient's preadmission residence and discharge destination and Medicare's Part B files of physician bills to study duration of nursing home stay. When a physician bills a nursing home for a visit, either the place of service is designated as a nursing home or a special visit code is used. This information was available for patients in three of the five sampled states. In states A and B, we report the number of patients for whom a bill was submitted during months 5, 6, or 7 after hospital admission for a physician visit to a skilled nursing home or other (residential) nursing home. In

state C, we report the number of patients for whom a bill was submitted for a visit to a skilled nursing facility.

To study hospital readmissions, we matched our patients to Health Care Financing Administration's bill retrieval file. Of the 92% of patients for whom we had accurate mortality data, we matched 96% for an overall success rate of 88%. We studied hospital readmission within 180 and 365 days postadmission and total days of acute care hospitalization. We included readmissions to all acute care hospitals regardless of the reason for readmission.

ADJUSTING OUTCOMES FOR SICKNESS AT ADMISSION

For length of stay and discharge destination (eg, home or nursing home), unadjusted and adjusted results are similar, and we present unadjusted data. For 30-day postadmission mortality, we adjusted pre- vs post-PPS differences using our 30-day disease-specific scale.[3-8] For 180-day postadmission mortality, prolonged nursing home stay, and hospital readmissions, we used the 30- and 180-day scales.[3-8] To compute, for example, pre- and post-PPS 30-day mortality adjusted for sickness at admission, we regressed 30-day mortality on a PPS indicator variable and the 30-day sickness scale and computed adjusted mortality rates pre- and post-PPS.

RESULTS

Length of Stay

For each of the diseases, length of stay fell: 21% for congestive heart failure, 18% for acute myocardial infarction, 14% for pneumonia, 32% for cerebrovascular accident, and 28% for hip fracture (Table 1). Overall, we found a 24% reduction in length of stay (from 14.4 to 11.0 days; 95% confidence interval, 3.1 to 3.8).

Mortality

The adjusted in-hospital mortality dropped from 16.1% to 12.8% (Table 2). Unweighted (for our sample design) adjusted mortality rates 30 days after admission for the five diseases combined were 16.7% pre- and 15.7% post-PPS, a difference of 1 percentage point (95% confidence interval, −0.1 to 2.1; $P=.07$). After reweighting our sample to represent the nation, values were 16.5% pre- vs 15.4% post-PPS, a difference of 1.1 percentage points ($P=.04$; Table 2).

As of 180 days posthospital admission, the adjusted mortality rates were 29.6% pre- and 29.0% post-PPS. Thus, almost one third of Medicare patients hospitalized with our five study diseases died within 6 months after admission. For congestive heart failure, cerebrovascular accident, and hip fracture, 180-day mortality dropped (significantly for hip fracture: 17.9% pre- and 14.8% post-PPS; $P<.05$), while for acute myocardial infarction and pneumonia, mortality rose post-PPS ($P>.05$). The pre- and post-PPS survival curves are similar for all conditions (Figure).

Discharge Destination

For the five diseases combined, the fraction of patients with a preadmission residence of home and a discharge destination of home was 77% pre- and 73% post-PPS ($P<.05$), with the most important difference being for hip fracture (56% pre- and 48% post-PPS, $P<.05$; Table 3). Overall, 95% of patients admitted from a nursing home returned to a nursing home, and this did not vary significantly by disease or time period.

Prolonged Nursing Home Stay

In studying prolonged nursing home stay, we focused on patients whose preadmission residence was home and who were still alive 7 months after the initial hospitalization. For the five study diseases combined, 8% of such patients in states A and B were in some type of nursing home approximately 6 months after the acute hospitalization, while 2% of such patients in state C were in a skilled nursing home (Table 4). In all three states, more, but not significantly more, patients had prolonged nursing home stays during the post-

Table 1.—Length of Stay, by Disease, Before and After Introduction of the Prospective Payment System (PPS)

Disease	n	Mean Length of Stay, d*† Pre-PPS (1981-1982)	Post-PPS (1985-1986)	Difference
Congestive heart failure	2824	11.1	8.8	−2.3‡
Acute myocardial infarction	2853	12.7	10.4	−2.3‡
Pneumonia	2749	12.1	10.4	−1.7‡
Cerebrovascular accident	2824	16.2	11.1	−5.1‡
Hip fracture	2762	20.1	14.5	−5.6‡
5 Diseases	14 012	14.4	11.0	−3.4‡

*Results were unadjusted for sickness at admission (unadjusted and adjusted results were similar).
†If the patient was discharged to a nonacute or "swing" hospital bed, the patient was considered discharged.
‡$P<.001$ for comparison of length of stay pre- vs post-PPS.

Table 2.—Adjusted Mortality Rates Before and After Prospective Payment System (PPS), by Disease and Type of Mortality Measure*

Mortality Adjusted for Sickness at Admission	Mortality Rates, % Dead† CHF Pre-PPS (n=1359)	CHF Post-PPS (n=1465)	AMI Pre-PPS (n=1416)	AMI Post-PPS (n=1437)	PNE Pre-PPS (n=1341)	PNE Post-PPS (n=1408)	CVA Pre-PPS (n=1382)	CVA Post-PPS (n=1442)	HIP Pre-PPS (n=1358)	HIP Post-PPS (n=1404)	5 Diseases Pre-PPS (n=6856)	5 Diseases Post-PPS (n=7156)
In-hospital mortality	12.3	8.9‡	24.0	21.8§	15.5	12.6‖	22.4	17.8‡	5.7	3.3‡	16.1	12.8‡
30-day postadmission mortality	14.7	13.0	24.2	24.2	15.9	15.7	21.3	19.9	5.3	4.6	16.7	15.7§¶
180-day postadmission mortality	33.5	31.7	33.6	34.8	27.8	29.2	35.3	34.3	17.9	14.8‖	29.6	29.0

*CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.
†In-hospital and 30-day postadmission mortality rates are adjusted for sickness at hospital admission using scales designed to predict death at 30 days postadmission; 180-day postadmission mortality rates are adjusted using scales designed to predict death at 180 days postadmission.
‡$P<.01$.
§$P=.05$ to .09.
‖$P<.05$.
¶As noted in the "Methods" section of the text, these data are unweighted. Reweighting for national representativeness changes the 30-day postadmission mortality rates for the five diseases combined as follows: 16.5% pre-PPS and 15.4% post-PPS ($P=.04$). This is the only outcome comparison whose significance is affected by the reweighting.

PPS period. The average increase across the three states was 1 percentage point ($P>.05$; 95% upper confidence bound, 2.2 percentage points).

Readmissions

As of 180 days after admission, the number of patients who died or who had at least one hospital readmission was unchanged: 57% pre- and 56% post-PPS ($P>.05$; Table 5). Results varied slightly by disease, with fewer patients with congestive heart failure, pneumonia, and hip fracture post-PPS having either a death or readmission ($P>.05$), but more patients with acute myocardial infarction suffering one of these two outcomes post-PPS ($P<.05$).

As of 365 days postadmission, the proportion of those patients discharged alive who had at least one hospital readmission was lower post-PPS for all diseases except acute myocardial infarction, and lower for the five diseases combined ($P<.05$ for congestive heart failure and hip fracture, and $P>.05$ for the other diseases individually and overall). Across all diseases except acute myocardial infarction, the total number of days spent in the hospital within 1 year of the study hospitalization was significantly lower post-PPS than pre-PPS ($P<.05$ for congestive heart failure, pneumonia, hip fracture, and the five diseases combined).

Summarizing Comparisons of Outcomes Pre- and Post-PPS

For the five diseases combined, in-hospital mortality was 3 percentage points lower post- vs pre-PPS ($P<.01$). However, this post-PPS improvement in mortality decreased to 1.1 percentage points by 30 days postadmission and to 0.6 percentage point by 180 days postadmission (Table 6).

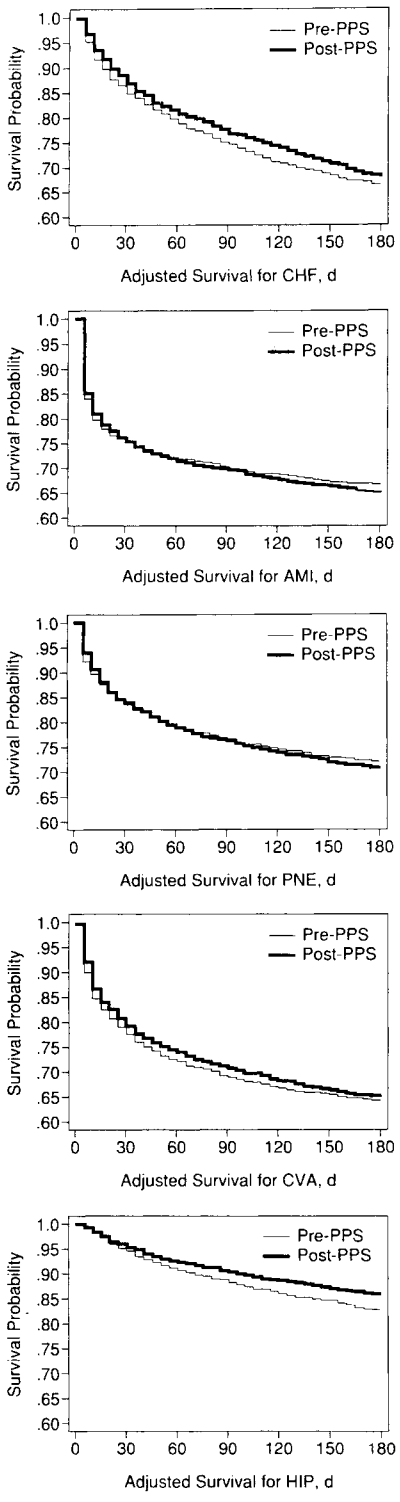
For patients admitted to the hospital from home, for the five diseases combined, we found that 4% more patients post-PPS were not discharged home ($P<.05$; 95% confidence interval for difference, 2.3 to 5.7 percentage points). We found that an additional 1% of patients (0% to 3% depending on the state) with a preadmission residence of home had evidence following hospitalization of a prolonged nursing home stay post-PPS ($P>.05$). For the five diseases combined, the proportion of patients with one or more hospital readmissions within 1 year of the initial hospitalization was 2 percentage points lower post-PPS than pre-PPS ($P>.05$).

COMMENT

Before discussing the relationship between the introduction of the PPS and changes in medical outcomes, it is important to note the severe burden of illness carried by elderly patients hospitalized with one of our five study conditions. By 1 month posthospital admission, 16% of these patients had died, with the death rate climbing to 29% within 6 months of admission. For hip fracture, the 6-month mortality rate is 16%, but for our other four medical conditions, it is over 33%. Of those patients who survived the initial hospitalization, more than half were readmitted in the next year. This fraction is highest for patients initially hospitalized with congestive heart failure (66%) and lowest for patients with hip fracture (44%). In addition, 25% of patients admitted from home and discharged alive are discharged to an institution. Forty-one percent of patients with cerebrovascular accident and 52% of patients with hip fracture previously living at home were discharged to an institution. However, most such institutional stays are short: of the patients still alive 7 months after hospital admission, 6% were in a nursing home.

Prior to and since the implementation of the PPS, clinicians, patients, and families have feared, and in some instances have reported, disasters in outcomes of care that were thought to be related to the new financial incentives. We have measured outcomes pre- and post-PPS on a nationally representative sample of more than 14 000 patients who were hospitalized with one of five diseases that make up 19% of Medicare admissions and 32% of deaths within 30 days. In contrast to these fears and anecdotal reports, we find no significant changes for the worse in either mortality at 30 and 180 days posthospital admission on the one hand or readmission and prolonged nursing home stay on the other.

We did find a significant increase in the fraction of patients discharged directly to an institution, but this does not appear to have resulted in a significant increase in prolonged nursing home stay. These last results, based on data from about 150 hospitals in three states, are not consistent with the findings of Fitzgerald et al[9,10] of increased prolonged nursing home stays post- vs pre-PPS for hip fracture patients (their findings were developed from two large hospitals). Our findings are consistent, however, with the clinical study of Mayer-Oakes et al[11] that examined outcomes of intensive care unit patients in three hospitals, the study of DesHarnais et al[12] that used secondary data sets to analyze inpatient mortality, and the studies of Palmer et al[13] and Gerety et al[14] that examined outcomes for hip fracture patients.



Survival curves comparing mortality following hospital admission through 180 days pre- and post-prospective payment system (PPS) for five diseases. CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.

Table 3.—Discharge Destination for Patients, by Disease* and Preadmission Residence

Admitted From and Discharged to	Patients Discharged to Each Destination†											
	CHF		AMI		PNE		CVA		HIP		5 Diseases	
	Pre-PPS	Post-PPS	Pre-PPS	Post-PPS	Pre-PPS	Post-PPS	Pre-PPS	Post-PPS	Pre-PPS	Post-PPS	Pre-PPS	Post-PPS
Home, % (n)	90 (1007)	88 (1133)	87 (960)	83‡(990)	92 (853)	88‡(875)	60 (922)	58 (1005)	56 (894)	48‡(987)	77 (4636)	73‡(4990)
Nursing home, % (n)	93 (82)	97 (99)	94 (32)	92 (39)	94 (204)	94 (268)	91 (90)	95 (114)	97 (301)	98 (263)	95 (709)	96 (783)

*CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.
†The entries in this table represent the number of patients discharged to each destination divided by the number of patients discharged alive, among those admitted from the indicated origin.
‡$P<.05$.

Table 4.—Residence in Nursing Home 6 Months Following Hospital Admission, by State and Disease, Pre– and Post–Prospective Payment System (PPS)

Disease‡	Patients Alive at 7 mo Posthospitalization Who Lived in Nursing Homes, %*								
	State A†			State B†			State C†		
	n§	Pre-PPS	Post-PPS	n§	Pre-PPS	Post-PPS	n§	Pre-PPS	Post-PPS
CHF	269	5.1	4.6	260	5.0	5.0	260	0.8	2.2
AMI	280	1.4	2.9	267	1.6	0.0	269	0.0	2.0
PNE	240	4.8	7.8	236	3.4	4.3	259	1.7	2.2
CVA	275	10.9	18.1	265	19.0	18.1	250	6.0	4.5
HIP	302	10.7	12.4	263	10.7	12.0	271	2.6	1.9
5 Diseases	1366	6.7	9.2	1291	7.9	8.1	1309	2.2	2.5

*The entries in this table represent the number of patients with evidence for physician nursing home bills submitted within months 5, 6, or 7 after admission divided by the number of patients alive 7 months after admission. This analysis only includes patients with a preadmission residence of home. None of the differences were significant at $P<.05$.
†Each fiscal carrier providing data for this analysis offered the data in a different format. For states A and B, the table compares the percentages of patients with evidence for prolonged nursing home stay indicated by a physician's bill to a skilled nursing or other (residential) nursing home. For state C, the table compares the percentages with evidence for prolonged nursing home stay as indicated by a physician's bill to a skilled nursing facility only.
‡CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.
§Pre- and post-PPS sample combined.

Table 5.—Hospital Readmissions Pre– and Post–Prospective Payment System (PPS) Adjusted for Sickness at Admission, by Disease*

Adjusted Outcome†	Outcome Within Time After Admission, No. of d	CHF		AMI		PNE		CVA		HIP		5 Diseases	
		Pre-PPS (n = 1078)	Post-PPS (n = 1247)	Pre-PPS (n = 1055)	Post-PPS (n = 1144)	Pre-PPS (n = 960)	Post-PPS (n = 1123)	Pre-PPS (n = 1164)	Post-PPS (n = 1273)	Pre-PPS (n = 985)	Post-PPS (n = 1045)	Pre-PPS (n = 5242)	Post-PPS (n = 5832)
Patients with death or readmission within 180 days, %‡	180	68	65	60	64¶	52	50	63	63	42	39	57	56
Patients with readmission within 1 year, %§	365	69	65¶	56	60	52	48	57	57	48	42¶	56	54
Mean No. of in-hospital days within 1 year‖	365	18	13¶	11	11	12	9¶	14	13	11	8¶	13	11¶

*CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.
†Outcomes are adjusted for sickness at admission using scales designed to predict death at 30 and 180 days postadmission.
‡Death (in-hospital or postdischarge) or any readmission within 180 days after admission.
§Any readmission within 1 year of admission among patients discharged alive from the initial hospitalization.
‖Mean number of in-hospital days within 365 days after the initial hospitalization for all sampled patients discharged alive regardless of readmission.
¶$P<.05$ for comparison of outcomes pre- and post-PPS.

Table 6.—Summary of Changes in Outcome Rates Post– vs Pre–Prospective Payment System (PPS)

Outcome Rates	Amount of Change (Post-PPS – Pre-PPS), %*					
	CHF	AMI	PNE	CVA	HIP	5 Diseases
In-hospital mortality†	−3#	−2#	−3#	−5#	−2#	−3#
30-day postadmission mortality†	−2	0	0	−1	−1	−1#
180-day postadmission mortality†	−2	+1	+1	−1	−3	−1
Patients admitted from home not discharged home‡	+2	+4#	+4#	+2	+8#	+4#
Prolonged nursing home stay§	0	+1	+2	+2	+1	+1
180-day postadmission mortality or readmission†‖	−3	+4#	−2	0	−3	−1
365-day postadmission readmission†¶	−4#	+4	−4	0	−6#	−2

*CHF indicates congestive heart failure; AMI, acute myocardial infarction; PNE, pneumonia; CVA, cerebrovascular accident; and HIP, hip fracture.
†Adjusted for sickness at admission.
‡Only includes patients admitted to hospital from home.
§Prolonged nursing home stay is measured by a physician's bill to a skilled nursing facility or other (residential) nursing home within months 5, 6, and 7 postadmission for three states combined, but only for patients admitted to hospital from home and alive as of 7 months postadmission.
‖Death or readmission within 180 days postadmission among all sampled patients.
¶Readmission within 1 year postadmission among all sampled patients.
#$P<.05$.

We were unable to prove that the PPS did not have a negative impact on outcome. Because of the manner in which the PPS was introduced, it was not possible to do a controlled trial of its effects, which would have permitted comparison of actual outcomes in the late 1980s under the PPS with outcomes during the same time period in the absence of the PPS. Valid causal conclusions from the only available data, which are observational in character, are difficult to achieve; we explore these issues in another article in this series.[15]

Moreover, our post-PPS data came from 1985 and 1986, a period in which the financial bases for prospective payment were still changing. The rates paid to hospitals at that time still included a component based on historical charges

during the cost-plus reimbursement era prior to the introduction of the PPS. In addition, the amount of reimbursement for a patient with a given diagnosis has been changing since 1986, and hospitals' financial margins have diminished. Evidence from another study, reported in this issue, indicates that clinical instability at discharge has increased post-PPS.[16] Last, although on average outcomes have not worsened, it is possible that some groups of patients may have suffered (eg, the old). Analyses addressing this last issue are ongoing.

In sum, the PPS was implemented in a manner that did not adversely affect overall outcomes for hospitalized Medicare patients. However, further clinical monitoring on a national level of the impact of the PPS is indicated to ensure that the continuing changes in the PPS do not negatively affect the outcomes of care.

This study was supported by cooperative agreement 18-C-98853/9-03 issued by the Health Care Financing Administration of the US Department of Health and Human Services, Baltimore, Md. The entire RAND costs ($3.9 million) of this project were funded with federal money.

We gratefully acknowledge the collaborative effort of the five professional review organizations that enabled this work to be completed. Participants from these organizations included medical directors, physician consultants and reviewers, project directors, and review coordinators from each of the five study states. In particular, we appreciate the keen clinical insights of the professional review organizations' physician specialists with whom we consulted throughout this study. We also acknowledge the many contributions of Harry Savitt, PhD, project officer from the Office of Research and Demonstrations of the Health Care Financing Administration of the US Department of Health and Human Services, whose administrative skills, astute commentary, and continuing support helped us immeasurably. Maureen Carney, MS, and Patricia A. Damiano were key contributors to project development and implementation. Andrea Steiner, MS, gave valuable editorial advice. Our policy advisory board (John C. Beck, MD, Barbara J. Burns, PhD, Monroe T. Gilmour, MD, Paul F. Griner, MD, Charlene Harrington, RN, PhD, T. Reginald Harris, MD, Rosalie Kane, DSW, Shirley Kellie, MD, Judith R. Lave, PhD, Charles E. Lewis, MD, Joseph Martin, Francis D. Moore, Jr, MD, Richard N. Pierson, Jr, MD, and James F. Rodgers, PhD) provided astute advice and guidance. We acknowledge the important contributions of the 297 hospitals whose medical records we reviewed. Without the efforts of these individuals and institutions, this evaluation could not have been successfully completed. Finally, we recognize and thank Florence McGinty, without whose secretarial skills this article would not have been produced.

References

1. Kahn KL, Brook RH, Draper D, et al. Interpreting hospital mortality data: how can we proceed? *JAMA.* 1988;260:3625-3628.
2. Draper D, Kahn KL, Reinisch E, 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.
3. Keeler E, Kahn KL, Draper D, et al. Changes in sickness at admission following the introduction of the prospective payment system. *JAMA.* 1990;264:1962-1968.
4. 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 Corp; 1988.
5. 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 Corp; 1988.
6. 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 Corp; 1988.
7. 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 Corp; 1988.
8. Roth CP, Kahn KL, Sherwood MJ, et al. *Medical Record Abstraction Form and Guidelines for Assessing Quality of Care for Hospitalized Patients With Pneumonia.* Santa Monica, Calif: Rand Corp; 1988.
9. Fitzgerald JF, Fagan LF, Tierney WM, et al. Changing patterns of hip fracture care before and after implementation of the prospective payment system. *JAMA.* 1987;258:218-221.
10. Fitzgerald JF, Moore PS, Dittus RS. The care of elderly patients with hip fracture. *N Engl J Med.* 1988;319:1392-1397.
11. Mayer-Oakes SA, Oye RK, Leake B, et al. The early effect of Medicare's prospective payment system on the use of medical intensive care services in three community hospitals. *JAMA.* 1988;260:3146-3149.
12. DesHarnais S, Chesney J, Fleming S. Trends and regional variations in hospital utilization and quality during the first two years of the prospective payment system. *Inquiry.* 1988;25:374-382.
13. Palmer RM, Saywell RM Jr, Zollinger TW, et al. The impact of the prospective payment system on the treatment of hip fractures in the elderly. *Arch Intern Med.* 1989;149:2237-2241.
14. Gerety MB, Soderholm-Difatte V, Winograd CH. Impact of prospective payment and discharge location on the outcome of hip fracture. *Gen Intern Med.* 1989;4:388-391.
15. Rogers WH, Draper D, Kahn KL, et al. Quality of care before and after implementation of the DRG-based prospective payment system: a summary of effects. *JAMA.* 1990;264:1989-1994.
16. Kosecoff J, Kahn KL, Rogers WH, et al. Prospective payment system and impairment at discharge: the 'quicker-and-sicker' story revisited. *JAMA.* 1990;264:1980-1983.

Quality of Care Before and After Implementation of the DRG-Based Prospective Payment System

A Summary of Effects

William H. Rogers, PhD; David Draper, PhD; Katherine L. Kahn, MD; Emmett B. Keeler, PhD; Lisa V. Rubenstein, MD, MSPH; Jacqueline Kosecoff, PhD; Robert H. Brook, MD, ScD

In this series we have described changes in the quality of care that have occurred in the treatment of hospitalized elderly Medicare patients with one of five conditions between 1981-1982 and 1985-1986. In this article we report on a mortality analysis, patient and hospital subgroup comparisons, and time series studies we have conducted in an attempt to determine whether changes in quality of care can be linked causally to the introduction of the prospective payment system. Based on these analyses we conclude that (1) mortality following hospitalization has been unaffected by the introduction of the prospective payment system, and improvements in in-hospital processes of care that began prior to the prospective payment system have continued after its introduction, but (2) the prospective payment system has increased the likelihood that a patient will be discharged home in an unstable condition. We recommend that efforts to correct this problem be intensified and that clinical monitoring of the impact of the prospective payment system continue as hospital cost-containment pressures intensify.

(*JAMA*. 1990;264:1989-1994)

IN 1984, the Health Care Financing Administration made a remarkable change in the Medicare system for financing hospital care for the elderly in the United States. Instead of paying for hospitalization on a cost basis, the Health Care Financing Administration developed a fixed-fee prospective payment system (PPS) based on diagnosis related groups. Because the new payment system contains incentives to reduce length of stay and substitute lower-cost services for more expensive ones, concern has arisen that quality may have declined under the PPS.[1-6]

We have conducted a nationally representative study of the effects of the PPS on the quality of care given to hospitalized Medicare patients.[7-13] Our study had two purposes: (1) to evaluate the quality of care given to the nation's Medicare patients before and after the introduction of the PPS and (2) to compare observed quality after the introduction of the PPS with predictions of what quality might have been in the same period in the absence of the PPS intervention. In previous articles in this series we have documented significant differences in sickness at admission,[9] processes of care,[10] unstable condition at discharge,[12] and outcomes[13] for Medicare patients hospitalized before and after the implementation of the PPS. Specifically, we demonstrated that, during the 1985-1986 study period, after the introduction of the PPS, the incidence of sickness at admission was higher, in-hospital processes of care were better, the number of patients discharged in unstable condition was higher, and mortality rates both 30 and 180 days following admission were lower or unchanged compared with 1981-1982, before the introduction of the PPS. In this report we attempt to sort out which of these differences may have been caused by PPS and which by other changes that occurred during the same period.[14]

For instance, in the last 10 years the number of patients treated in outpatient settings has risen and the average burden of illness of such patients has also increased. The use of do-not-resuscitate (DNR) orders has grown.[15,16] Professional review organizations were established and have been extending their reach and insisting on better accountability. New technologies and medical knowledge have become available, and as time passes older physicians are continually being replaced by younger physicians who have more thorough train-

From the Health Program of the RAND Corp, Santa Monica, Calif (Drs Rogers, Draper, Kahn, Keeler, Rubenstein, and Brook); the Departments of Medicine (Drs Kahn, Rubenstein, Kosecoff, and Brook) and Health Services (Drs Kosecoff and Brook), UCLA; and Value Health Sciences Inc, Santa Monica (Dr Kosecoff)

The opinions, conclusions, and proposals in the text are those of the authors alone and do not necessarily represent the views of the RAND Corp or UCLA.

Reprint requests to the RAND Corp, 1700 Main St, PO Box 2138, Santa Monica, CA 90406-2138 (Dr Rogers).

ing in the use of newly developed medical technology.

We report on three types of analyses that supplement our previous descriptive before-and-after comparisons and help us determine which of the observed changes were caused by the PPS. First, we estimate the impact on mortality of each of the previously described changes associated with the PPS. Second, for the important changes, we examine patterns of change among patient subsets to see whether the changes are consistent across all patient groups or are concentrated in particular types of patients and hospitals. Third, we extend our before-and-after comparisons to a time series analysis with multiple points before and after introduction of the PPS and examine trends within and across the 1981-1982 and 1985-1986 periods to determine whether values after the introduction of the PPS are consistent with trends before the PPS.

Table 1.—Variables Used to Study Effects of Prospective Payment on Mortality Following Hospitalization*

Age.—Age in years.
Gender.—Female (scored 1) vs male (scored 0).
Race.—Black (scored 1) vs other (scored 0).
Medicaid.—Patients with Medicaid (scored 1) vs all other patients (scored 0).
Preadmission Nursing Home Residence.—Patients who resided in a nursing home immediately before the study hospitalization (scored 1) vs all other patients (scored 0). A *nursing home* was defined as a skilled nursing facility, intermediate-care facility, extended-care facility, or an unspecified nursing home.
30-Day Mortality Prediction.—A disease-specific weighted sum of variables selected by logistic regression of 30-day mortality rates on clinical conditions present at the time of the study hospitalization (see Keeler et al[9] for details).
180-Day Mortality Prediction.—A disease-specific weighted sum of variables selected by logistic regression of 180-day mortality rates on clinical conditions present prior to admission for the study hospitalization (see Keeler et al[9] for details).
Initial DNR Orders.—Patients for whom do-not-resuscitate (DNR) order was written by the physician on day 1 or 2 of the study hospitalization (scored 1) vs all other patients (scored 0).
Process of Care.—Five standardized process scales summarizing disease-specific process measures, including physician cognitive diagnostic processes, nurse cognitive diagnostic processes, technical diagnostic processes, technical therapeutic processes, and intensive care unit/telemetry monitoring, plus an overall summary scale (see Kahn et al[10] for details).
Subsequent DNR Orders.—Patients for whom DNR order was written by the physician on day 3 of the study hospitalization or later (scored 1) vs all other patients (scored 0).
Unstable Condition at Discharge.—Patients with at least one clinical problem noted at discharge (1) that should have been corrected prior to discharge and (2) that was likely to be associated with worse patient outcomes (scored 1) vs all other patients (scored 0) (see Kosecoff et al[12] for details).
Abnormal Last Laboratory Values.—Patients with at least one dangerously abnormal laboratory value the last time laboratory values were recorded during the study hospitalization (scored 1) vs all other patients (scored 0) (see Kosecoff et al[12] for details).
Sickness at Discharge.—Patients with either unstable condition at discharge or another measure of clinical sickness at the time of hospital discharge (scored 1) vs all other patients (scored 0) (see Kosecoff et al[12] for details).
Discharge Destination.—Patients discharged to an institution, including nursing homes (skilled, intermediate-care, or extensive-care facilities or unspecified), retirement homes (sheltered housing, congregate housing, halfway houses, or board-and-care facilities), psychiatric facilities (mental hospitals), chronic disease hospitals, or rehabilitation hospitals (scored 1) vs all other patients (scored 0) (see Kahn et al[13] for details).

*All variables were abstracted from the medical record.

METHODS

Sampling

The sampling design is described elsewhere in this series.[8]

Estimating the Impact on Mortality of Changes After the Introduction of the PPS

We identified 14 variables, each of which is associated with changes in mortality 30 and 180 days after admission (Table 1). We conducted a mortality analysis in which we used linear regression to estimate how changes in death rates between 1981-1982 and 1985-1986 were associated with each of the 14 variables.[17] We then postulated a sequence of effects corresponding to the temporal order in which the 14 variables potentially affect the patient. In this temporal order, the five demographic variables came first (introduced simultaneously into the multivariate analysis), followed by a 30- and 180-day sickness-at-admission scale (the admission sickness measures most highly correlated with death by 180 days after admission); initial DNR status; a summary of in-hospital processes of care; subsequent DNR status; and four discharge status variables: unstable condition at discharge, abnormal last laboratory value, sickness at discharge, and discharge destination.

We then computed, using the method of indirect effects,[18] the expected change in 30- and 180-day mortality rates associated with the change in each variable, adjusting for all earlier variables in the temporal sequence. These analyses were designed to tell us which variables potentially had the most influence on death. Such analyses can be used to demonstrate association but not necessarily causation. Sequencing the variables in an order different from the one we chose would produce different effect estimates for the variables (except for demographics, which were entered simultaneously).

We also tested for the ability to predict death of interactions between the variables and the experimental variable before vs after the introduction of the PPS. If present, such interactions would signify changes in the relationship between the variables and death in the two periods. A standard F test was used.[19]

Patient and Hospital Subset Analyses

In our second analysis we examined whether differences after the introduction of the PPS occurred globally across all patient and hospital subsets or were restricted to certain patient or hospital groups. Underlying this analysis was the idea that a large systemic change, such as Medicare's movement from retrospective to prospective payment, should have had pervasive effects that appear consistently across a wide variety of patient and hospital types. Less global changes, such as diffusion of current standards of medical practice from urban to rural hospitals, that took place at about the same time as the introduction of the PPS might have had isolated effects that were concentrated in certain classes of patients and hospitals.

We examined how the process of in-hospital care, initial use of DNR orders, and discharge status changed across patient subsets. The patient subsets differed by levels of sickness and functioning, and the hospital subsets differed by urbanicity, state, percentage of Medicaid patients, and teaching status. We computed separately for each of five diseases the changes after the introduction of the PPS in initial use of DNR orders, process of care, and unstable condition at discharge for each of these subsets. Next, we developed a hospital-level variable that combined urbanicity and teaching as follows: (1) rural nonteaching, (2) urban nonteaching, and (3) urban teaching, because these were the most interesting distinctions.[17] We used linear regression to assess the significance of differences in process, initial use of DNR orders, and discharge status for patients in these three subsets.

Trend Analyses

Finally, we made graphic comparisons between (1) the actual values of important quality-of-care variables in 1985-1986 and (2) simple estimates of what those values might have been if the PPS had not been implemented. To do this we used our patient-level data to calculate quarterly summaries of four variables: 180-day mortality, adjusted for sickness at admission; overall in-hospital process of care; unstable condition at discharge; and initial use of DNR or-

ders. We plotted these quarterly values on the same graphs as predictions based on simple extrapolations from the period before the PPS was introduced. With only eight quarters of data before the PPS period was introduced and with the PPS introduced nonexperimentally, our ability is limited to definitively say what the quality of care would have been like in the period after the introduction of the PPS in the absence of the PPS (for example, the confidence bands based on linear extrapolation models are quite wide). However, if changes between 1981-1982 and 1985-1986 are consistent with trends within those two periods, it would appear that the changes are a continuation of trends in place historically and not a specific result of the introduction of the PPS.

RESULTS

Impact on Mortality of Changes After the Introduction of the PPS

The major effects on 180-day mortality associated with changes after the introduction of the PPS are concentrated in four variables: sickness at admission, initial DNR status, in-hospital process of care, and discharge status. See Table 2 for effects after adjustment for temporal sequencing. The effects on 30-day mortality rates are very similar, so they are not shown. Increases in sickness at admission, adjusting for changes in demographics, had an especially pronounced predicted effect on mortality for patients with pneumonia (an expected rise of 3.0 percentage points in the 180-day mortality rate) and hip fracture. After adjusting for changes in demographics and sickness at admission, the rise in the initial use of DNR orders was associated with a modest increase (0.2 to 0.7 percentage points) in expected mortality for all five diseases.

Improvements in in-hospital process of care after the introduction of the PPS (after accounting for changes in demographics, sickness at admission, and initial DNR orders) were associated with noticeable decreases in expected mortality for all five diseases (0.5 to 1.3 percentage points). These changes were significant for all conditions but hip fracture. There was a statistically significant rise in expected mortality associated with adverse changes in one or more discharge variables in four of the five diseases. Changes in the number of patients in unstable condition at discharge and the number of patients with abnormal last laboratory values were significant for three of the five diseases. Taken together, problems at discharge after the introduction of the PPS have increased enough to raise expected death rates by an amount that varies from a total of 0.2 percentage points for patients with acute myocardial infarction to 0.9 percentage points for patients with pneumonia.

Table 2.—Predicted Change in 180-Day Mortality Rates After the Introduction of the PPS*

Variable	Predicted Change in 180-Day Mortality Rates After the Introduction of the PPS, Percentage Points†				
	Congestive Heart Failure	Acute Myocardial Infarction	Pneumonia	Cerebrovascular Accident	Hip Fracture
Patient demographics‡	0.7§	1.2§	1.4§	0.3	−0.2
Age	0.1	0.8§	0.3	0.2	−0.1
Gender	0.2	0.0	0.0	0.0	0.3
Race	0.1	−0.0	0.1	0.0	0.0
Medicaid	0.0	0.0	0.0	0.0	0.0
Preadmission residence	0.3	0.4‖	1.0§	0.1	−0.4‖
Sickness at admission¶	−0.1	0.5	3.0§	−1.1	1.1§
Initial DNR orders	0.3	0.7§	0.5‖	0.5‖	0.2
Process of care	−1.0§	−0.7§	−1.3§	−1.0‖	−0.5
Subsequent DNR order	0.2	0.1	0.2	0.0	−0.0
Discharge problems#	0.6§	0.2	0.9§	0.6§	0.4
Unstable condition at discharge	0.2	0.0	0.2	0.5‖	0.1‖
Abnormal last laboratory value	−0.2	−0.0	0.2‖	−0.1	0.0
Sickness at discharge	0.4‖	0.2	0.3	0.1	−0.0
Discharge destination	0.2	−0.0	0.2‖	0.1	0.3‖
Total change accounted for**	0.7	2.0	4.7	−0.7	1.0
Actual Differences in observed death rates, per 100 patients††	−1.1	3.4	5.8	−1.2	−1.9
Remaining unexplained differences	−1.8	1.4	1.1	−0.5	−2.9

*PPS indicates prospective payment system; DNR, do not resuscitate.

†Except for patient demographics, the effect of each variable is the added effect, controlling for variables that appear higher in the table. The mortality effects displayed in this table are an average of the 1981-1982 period, when mortality effects were large, and the 1985-1986 period, when mortality effects were somewhat smaller.

‡The predicted changes for patient demographics are a sum of the predicted changes for age, gender, race, Medicaid, and preadmission residence.

§$P<.01$.

‖$P<.05$.

¶Sickness at admission is evaluated using two variables, a scale designed to predict death at 30 days and a scale designed to predict death at 180 days (see Table 1).

#The predicted changes for discharge problems are a sum of the predicted changes for unstable condition at discharge, abnormal last laboratory value, sickness at discharge, and discharge destination.

**Sum of the changes accounted for by the variables that appear higher in the table, subject to rounding error. For example, for congestive heart failure, the sum of the percentages listed above is 0.7%, meaning that mortality associated with the named factors is estimated to have increased by a total of 0.7 percentage points. The actual observed change in death rates is 1.1%, and the difference (−1.1 − 0.7) is due to other, unnamed causes.

††Actual 180-day mortality rates were as follows: for congestive heart failure, 33.0% before and 31.9% after the PPS; for acute myocardial infarction, 32.4% before and 35.9% after the PPS; for pneumonia, 25.6% before and 31.3% after the PPS; for cerebrovascular accident, 35.4% before and 34.2% after the PPS; and for hip fracture, 17.2% before and 15.3% after the PPS.

Generally, interactions between the effect of death on the variables in Table 2 and the period before or after the introduction of the PPS were statistically insignificant. For unstable condition at discharge, however, the differential in mortality rates associated with discharge in unstable condition in 1985-1986 was about 60% as large as the differential in mortality rates in 1981-1982. This interaction effect was insignificant for individual diseases but significant ($P<.05$) when combined across all five diseases.

At the bottom of Table 2 are the actual changes in death rates between 1981-1982 and 1985-1986, the total of all accounted-for expected changes, and the difference between these two figures. Our analysis explains only a portion of the actual differences. The discrepancies could be due either to sampling variability (the SEs for the changes in death rates are about 2 percentage points for each disease) or to systematic effects unaccounted for in our analysis of the sources of change. For example, our analysis does not address the contribution of new technology to lower death rates because we specifically excluded new technologies from our process measures.

Patient and Hospital Subset Analyses

Table 3 presents differences (after the PPS minus before the PPS) in the summary in-hospital process of care score (in SD units), in the percentage of patients with a DNR order on admission, and in the percentage of patients with one or more unstable conditions at discharge for each of the three levels of

Table 3.—Differences in Process of Care, Initial DNR Orders, and Unstable Condition at Discharge, by Urbanicity and Teaching Status of Hospitals After the Introduction of the PPS*

Characteristic	Change After the Introduction of the PPS†			P‡
	Rural Nonteaching	Urban Nonteaching	Urban Teaching	
Congestive Heart Failure (n=2812)				
Sample size	644	1182	986	...
Process of care, SD units	+0.73	+0.38	+0.27	<.01
Initial DNR orders, percentage points	+3.1	+0.0	+2.7	NS
Unstable condition at discharge, percentage points	+4.3	+3.5	+2.8	NS
Acute Myocardial Infarction (n=2841)				
Sample size	623	1238	980	...
Process of care, SD units	+0.40	+0.28	+0.17	<.05
Initial DNR orders, percentage points	+3.7	+3.0	+2.4	NS
Unstable condition at discharge, percentage points	+1.7	+1.1	+4.7	NS
Pneumonia (n=2727)				
Sample size	600	1169	958	...
Process of care, SD units	+0.56	+0.41	+0.39	NS
Initial DNR orders, percentage points	+3.5	+4.4	+3.5	NS
Unstable condition at discharge, percentage points	+11.6	−0.3	+2.5	<.10
Cerebrovascular Accident (n=2810)				
Sample size	601	1212	997	...
Process of care, SD units	+0.65	+0.49	+0.39	<.05
Initial DNR orders, percentage points	+1.4	+1.9	+4.0	NS
Unstable condition at discharge, percentage points	+6.3	+3.7	+2.3	NS
Hip Fracture (n=2746)				
Sample size	468	1282	996	...
Process of care, SD units	+0.58	+0.43	+0.33	<.01
Initial DNR orders, percentage points	−0.4	+0.91	+1.4	NS
Unstable condition at discharge, percentage points	+11.31	+3.1	+2.9	<.10

*PPS indicates prospective payment system.
†Entries are the "raw" differences after the introduction of the PPS. Adjustment for sickness at admission had little effect.
‡NS indicates not significant at the 10% level for differences after the introduction of the PPS among the three hospital types.

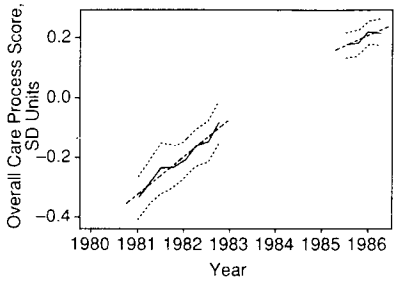


Fig 1.—In-hospital processes of care for five diseases (n=14 012). Solid line indicates values for the sample; dotted and dashed line, trends before and after the introduction of the prospective payment system; and dotted lines, sample values ±2 SEs.

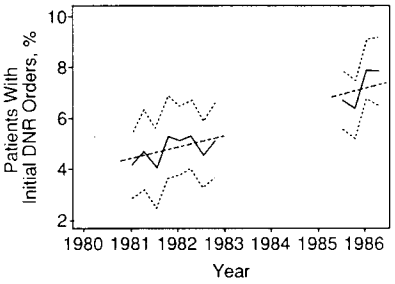


Fig 2.—Use of do-not-resuscitate (DNR) orders on day 1 or 2 of hospital stay for five diseases (n = 14 012). Solid line indicates values for the sample; dotted and dashed line, trends before and after the introduction of the prospective payment system; and dotted lines, sample values ±2 SEs.

the urbanicity/teaching hospital variable (rural nonteaching, urban nonteaching, and urban teaching). Across all five diseases we observed a consistent and significant pattern of differences by hospital setting in improvements in the process of care after the introduction of the PPS. Rural nonteaching hospitals showed the biggest gains in in-hospital process of care, and urban teaching facilities showed the smallest gains. In contrast, the initial use of DNR orders and the number of patients in unstable condition at discharge showed consistent increases across the three hospital types after the introduction of the PPS (Table 3).

We found some differences in the magnitude of the increase in use of DNR orders after the introduction of the PPS as a function of patient characteristics. The biggest increase was in patients with poor function and high levels of acute sickness at admission. Using the variables mentioned in the "Methods" section to define patient and hospital subgroups, we were not able to find any types of patients for any of our five study diseases for whom the rise in the number of patients in unstable condition at discharge after the introduction of the PPS differed significantly from the average increase in the number of patients in unstable condition for all patients with that disease.

Trend Analyses

Figure 1 shows that the quality of the in-hospital process of care was on the rise before the introduction of the PPS ($P<.01$) and continued to rise after the introduction of the PPS, although at a slightly slower rate (the difference between the rates is not significant). Figure 2 shows that use of initial DNR orders was increasing in 1981-1982 ($P<.05$) and that this increase continued in 1985-1986. The percentage of patients discharged in unstable condition was reasonably flat both in 1981-1982 and in 1985-1986, with a jump between these two periods (Fig 3). Sickness at discharge and discharge to a nursing home also had jumps between these periods (data not shown). Figure 4 presents quarterly values of 180-day mortality aggregated across five diseases, after adjusting for sickness at admission. There was no trend in either period, and the adjusted mortality values after the introduction of the PPS were consistent with values before the PPS.

COMMENT

After adjusting for sickness at admission, mortality was unchanged or lower in 1985-1986 after the introduction of the PPS than in 1981-1982. We believe this was the result of two counterbalancing forces: sharp improvements in in-hospital process of care from 1981 to 1986 that acted to lower mortality rates, offset by increases during the same period in sickness at admission,[9] the use of DNR orders,[15-16] and patients in unstable condition at discharge that acted to raise mortality.[12] However, most of the improvements in the process of care we have documented were probably not caused by the PPS. Two types of evidence presented herein support this conclusion: the presence of a significant upward trend in the process of care in 1981-1982, before the PPS was introduced, and the lack of uniformity of improvements in the care process across

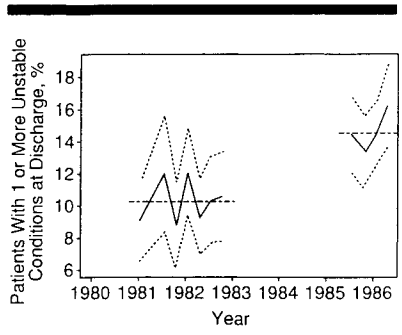


Fig 3.—Patients discharged with one or more unstable conditions for five diseases (n = 7412). Solid line indicates values for the sample; dotted and dashed line, trends before and after the introduction of the prospective payment system; and dotted lines, sample values ± 2 SEs.

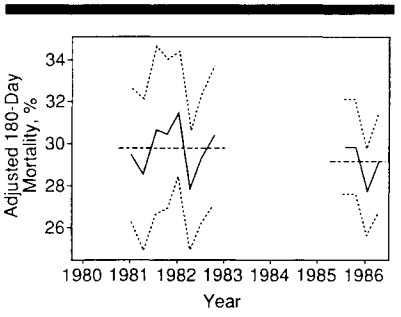


Fig 4.—Mortality rates 180 days after admission, adjusted for sickness at admission, for five diseases (n = 12 823). Solid line indicates values for the sample; dotted and dashed line, trends before and after the introduction of the prospective payment system; and dotted lines, sample values ± 2 SEs.

different types of patients and hospitals after the introduction of the PPS. It is difficult to understand why, if the PPS caused the gains in the process of care we have demonstrated, it did so with substantially more force in rural nonteaching facilities than in urban teaching hospitals. Rather, it seems more plausible that ongoing trends in medicine, such as diffusion of newer methods into outlying areas, were responsible for the observed improvement. On this score the PPS may be judged a success: the care process did not deteriorate in the effort to save money.

An increased rate of use of DNR orders at admission after the introduction of the PPS was associated with somewhat higher death rates. Do-not-resuscitate orders potentially save the hospital money and save the patients pain, but increased usage of DNR orders is associated with an increase in death rates. The patient subgroups in which the biggest changes in use of DNR orders were concentrated—those who were both acutely ill and functioning poorly—suggest that the potential benefits of medical care are being examined more closely for these patients. This makes sense and does not implicate the PPS.

The number of patients in unstable condition at discharge and related factors have risen in temporal association with the PPS. Furthermore, the number of patients in unstable condition at discharge is increasing across the board rather than in any specific patient or hospital subgroup and has increased in ways we believe are not due to changes in recording.[12] Both results implicate the PPS. On the face of it, this problem needs attention.

There is room, however, for substantial debate about how serious a problem this is. First, should *unstable condition* be defined broadly or narrowly? If unstable condition is defined narrowly to include only factors that are clearly remediable, the increase in the number of patients discharged in unstable condition after the introduction of the PPS may be responsible for a 0.5 percentage point increase in the mortality rate, depending on the disease. If unstable condition is defined broadly to include sickness at discharge and discharge to a nursing home, the associated rise in death rates may be as high as 0.9 percentage points. We do not know how fixable these problems are, but we do know that we measured only a fraction of the problems that should be fixed before discharge.

Second, there is no guarantee that the hospital would be able to reduce mortality by holding patients in unstable condition longer. Third, the effects of unstable condition might be reduced by more cost-effective mechanisms than increasing hospital length of stay. For example, patients in unstable condition could be given special consideration in nursing homes or could be part of well-designed posthospitalization home nursing programs. Indeed, the impact on mortality of discharge in unstable condition does not appear to be as great in 1985-1986 as it was in 1981-1982; this may be a result of improvements in care after hospitalization. Fourth, the length of stay has remained relatively constant since 1986, so any contribution of inappropriately short hospital stays to discharge problems has probably stabilized.[20]

On the other hand, entirely eliminating discharges of patients in unstable condition (not just the effect after the introduction of the PPS) might have a large impact on the effectiveness of hospital care. For the five diseases studied, patients in unstable condition at discharge after the introduction of the PPS have a mortality rate 30% higher than that of patients discharged in stable condition. This translates into additional mortality of 4.4 percentage points in the 90 days following admission for patients in unstable condition at discharge. An observational study such as ours cannot definitely estimate the effects of better discharge monitoring on patient outcomes; a controlled experiment with specific discharge protocols is required.

Between 1981-1982 and 1985-1986 there were also important changes in the demographics and sickness of patients. Patients with pneumonia are much sicker at admission now than before. For other diseases, the nature of the change in sickness at admission is less clear. As we mentioned above and have documented elsewhere,[9] some of the apparent increase in sickness appears to be a recording bias. When we move from recording-sensitive sickness indicators (such as the functional status and acute and chronic sickness measures in Table 1) to the recording-insensitive indicators used in the remainder of this study's analyses (such as our 30- and 180-day scales), the apparent rise in sickness at admission is smaller. Part of the rise in severity, however, may be due to efforts by the professional review organizations to restrict admissions of less-sick patients. For two diseases (acute myocardial infarction and pneumonia), admissions of the least-sick patients appear to be down about 30%.

In conclusion, because the PPS was not introduced as an experiment, our observational time series study can provide only limited answers about the changes in quality of care that the PPS, and the PPS alone, caused. As we have noted elsewhere in this series,[13] three other caveats are also worth bearing in mind when considering our results: (1) We studied only five diseases, albeit five important diseases in the Medicare cohort. (2) Our study design allowed us only to assess differences in quality of care once patients are hospitalized, so that, for example, we cannot comment on any changes in *access* to hospital care the PPS may have caused. (3) Our study covers the era after the introduction of the PPS only through June 1986. Since that time hospital payments under the PPS have been tightened.[7]

Even so, two key policy conclusions appear clear: (1) At least through the middle of 1986, the PPS did not interrupt an important long-term trend toward better processes of in-hospital care, a trend that has led to somewhat

lower death rates. (2) On the other hand, we believe that the PPS has had an adverse effect on the condition in which patients are discharged.

These conclusions lead to three major policy recommendations: (1) To eliminate any possible problems with patients in unstable condition at discharge, a more systematic assessment should be made of the readiness of a patient to leave the hospital and be cared for as an outpatient or in an institution other than an acute-care hospital. Perhaps our instability-at-discharge scale would be a good place to start. (2) If further investigation suggests that some discharges of patients in unstable condition may be acceptable with suitable follow-up or could be prevented by a longer hospital stay, we recommend that clinical trials be undertaken to evaluate the impact of such changes on mortality. (3) To provide current information about the effects of Medicare's payment methods on quality of care, we also recommend the continued collection of clinically detailed data that monitor sickness at admission, processes of care, discharge status, and outcomes on a regular basis as long as prospective payment is in place.

This study was supported by cooperative agreement 18-C-98853/9-03 issued by the Health Care Financing Administration of the US Department of Health and Human Services. The entire costs to the RAND Corp ($3.9 million) for this project were funded with federal money.

We gratefully acknowledge the collaborative effort of the five professional review organizations that enabled this work to be completed. Participants from the professional review organizations included medical directors, physician-consultants and -reviewers, project directors, and review coordinators from each of the five study states. In particular, we appreciate the keen clinical insight of the specialist physicians from the professional review organizations with whom we consulted throughout this study. We also acknowledge the many contributions of Harry Savitt, PhD, project officer from the Office of Research and Demonstrations of the Health Care Financing Administration of the US Department of Health and Human Services, whose administrative skills, astute commentary, and continuing support helped us immeasurably. Maureen Carney, MS; Caren Kamberg, MSPH; Ellen Reinisch, MS; and Marjorie J. Sherwood, MD, were key contributors to project development and implementation. Andrea Steiner, MS, gave valuable editorial advice. Our policy advisory board (John C. Beck, MD; Barbara J. Burns, PhD; Monroe T. Gilmour, MD; Paul F. Griner, MD; Charlene Harrington, RN, PhD; T. Reginald Harris, MD; Rosalie Kane, DSW; Shirley Kellie, MD; Judith R. Lave, PhD; Charles E. Lewis, MD; Joseph Martin; Francis D. Moore, Jr, MD; Richard N. Pierson, Jr, MD; and James F. Rodgers, PhD) provided astute advice and guidance. We acknowledge the important contributions of the 297 hospitals whose medical records we reviewed. Without the efforts of these individuals and institutions, this evaluation could not have been successfully completed. Finally, we thank Florence McGinty, without whose secretarial skills this article would not have been produced.

References

1. Omenn GS, Conrad DA. Implications of DRGs for clinicians. *N Engl J Med.* 1984;311:1314-1317.
2. Smits HL, Watson RE. DRGs and the future of surgical practice. *N Engl J Med.* 1984;311:1612-1615.
3. Stern RS, Epstein AM. Institutional responses to prospective payment based on diagnosis-related groups: implications for cost, quality, and access. *N Engl J Med.* 1985;312:621-627.
4. Iglehart JK. Early experience with prospective payment of hospitals. *N Engl J Med.* 1986;314:1460-1464.
5. McCarthy CM. DRGs: five years later. *N Engl J Med.* 1988;318:1683-1686.
6. Schramm CJ. Prospective payment: some retrospective observations. *N Engl J Med.* 1988;318:1681-1683.
7. 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.
8. 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.
9. 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.
10. 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.
11. Rubenstein LV, Kahn KL, Reinisch EJ, et al. Changes in quality of care for five diseases measured by implicit review, 1981 to 1986. *JAMA.* 1990;264:1974-1979.
12. Kosecoff J, Kahn KL, Rogers WH, et al. Prospective payment system and impairment at discharge: the 'quicker-and-sicker' story revisited. *JAMA.* 1990;264:1980-1983.
13. 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.
14. Vladeck BC. Hospital prospective payment and the quality of care. *N Engl J Med.* 1988;319:1411-1413.
15. Scheidermayer DL. The decision to forgo CPR in the elderly patient. *JAMA.* 1988;260:2096-2097.
16. Bedell SE, Delbanco TL. Choices about cardiopulmonary resuscitation in the hospital: when do physicians talk with patients? *N Engl J Med.* 1984;310:1089-1093.
17. Kahn KL, Draper D, Keeler EB, et al. *The Effects of the DRG-Based Prospective Payment System on Quality of Care for Hospitalized Medicare Patients: Design, Methods, and Results.* Santa Monica, Calif: RAND Corp; 1990:2. R-3931-HCFA.
18. Wonnacott TH, Wonnacott RJ. *Introductory Statistics for Business and Economics.* 3rd ed. New York, NY: John Wiley & Sons Inc; 1984:377-379, 686.
19. Scheffe H. *The Analysis of Variance.* New York, NY: John Wiley & Sons Inc; 1959.
20. Witsberger C, Kominski G. *Recent Trends in Length of Stay for Medicare Surgical Patients.* Santa Monica, Calif: RAND Corp; 1990. R-3940-HCFA.