Measuring The Quality Of Nursing Surveillance Activities For Five Diseases Before And After Implementation Of The Drg-based Prospective Payment System

Lisa V. Rubenstein, MD,MSPH
Betty L. Chang, RN,DNSc, FAAN
Emmett B. Keeler, PhD
Katherine L. Kahn, MD

As part of a national study on the effects of the Medicare Prospective Payment System (PPS) based on Diagnosis Related Groups (DRG's), RAND researchers collaborated with physicians from five state Peer Review Organizations (PRO's) to develop quality of care measures for five major medical conditions: acute myocardial infarction (AMI), acute cerebrovascular accident (CVA), pneumonia (PNE), congestive heart failure (CHF), and hip fracture (HIP). The quality of care measures used medical record review to determine the quality of inpatient care for Medicare patients age 65 and older admitted to an acute hospital with one of the five study diseases. The quality of care measures we used included scales focussing on physician care as well as scales focussing on nursing surveillance of key clinical signs and symptoms for each disease. The purpose of this paper is to establish the reliability and validity of our nursing surveillance scales, and to determine whether nursing surveillance, as evaluated by our measures, has changed since the advent of PPS.

METHODS

Patient Sample

The patient sample for this study includes 14,012 Medicare patients age 65 years and older from 297 hospitals. The 297 sampled hospitals represent 97% of the 305 hospitals approached regarding participation in the study. The hospitals were selected from 30 cities or towns in five states, each from a different region of the United States. Hospitals were chosen to represent the national patient cohort with respect to urbanicity, percentage of Medicare patients, hospital size, teaching intensity, and type of ownership. We oversampled hospitals caring for poor patients. Estimates weighted to reflect our oversampling differ little from unweighted estimates in our major analyses; we therefore present unweighted results (Draper D, Kahn KL, et. al., l990).

Patients hospitalized with congestive heart failure, acute myocardial infarction, pneumonia, cerebrovascular accident, and hip fracture were randomly selected from lists of all Medicare patients hospitalized during the study years. The study sampled patients from before PPS (1981 (20%), 1982 (30%), and after PPS (1985/1986 (50%). Lists of patients were based on ICD-9-CM codes, but patient medical records were then screened to exclude patients who did not have the study disease as the reason for admission (Draper D, Kahn KL, et. al., 1990).

Developing Process Criteria

We sought to develop process measures that were both clinically valid and reliably measured by retrospective medical record review (Kahn KL, Rogers WH, et. al, 1990). We first reviewed the literature for each study disease and consulted with experts, searching for those clinical processes that were most critical in determining patient outcomes. These measures were then presented

to physician panels consisting of 5 - 12 generalist and specialist physicians, who were selected by our collaborators in this study, the Professional Review Organizations (PRO's) from the five states in which the study's sampling occurred. Each panel reviewed the suggested criteria to decide whether they believed that the criteria were likely to be reliably recorded in the medical record in 1981 as well as in 1986, and whether the criteria made clinical sense. The study was initially designed as a physician study on quality of care. Our physician panelists unanimously agreed, however, that the quality of nursing care was a strong determinant of patient outcome. Process criteria measuring nursing surveillance were therefore included in our measures.

Some criteria were applicable to all patients, and some to subsets of patients with specific signs, symptoms, or characteristics. For example, all patients should be assessed to determine their ambulatory status prior to admission, but only patients with acute chest pain require sublingual nitroglycerin.

We then pilot-tested the validated criteria to determine whether they could, indeed, be reliably measured. For example, we found on testing that the number of times patients were out of bed was not easily measured from the record; this had been predicted by our panel. Similarly, neither the adequacy of nutrition nor the performance of patient education could be accurately assessed within the limits of our study design. On the other hand, the presence or absence of a nursing note about nutritional intake in stroke patients, and the presence or absence of a nursing note indicating whether patients with congestive heart failure or pneumonia were short of breath, could be measured.

Collecting Process Data

A multidisciplinary team, including physicians, social scientists, and a nurse with a master's degree in public health, then developed disease-specific abstraction forms to collect the medical record data (Kahn KL, Chassin MR et al., 1988; Roth CP, Kahn KL, et al., 1988; Rubenstein LV, Kosecoff J et al., 1988; Sherwood MJ, Kahn KL, et al., 1988; Kosecoff J, Kahn KL, et al., 1988). A doctoral degree prepared nurse joined the team for data analysis of the nursing surveillance measures.

In addition to studying some process measures that were applicable to all diseases, we studied some process criteria that were specific for each disease. We collected data on approximately 100 process criteria for each disease. Of these, approximately one-fifth measured nursing surveillance functions. Although our emphasis here is on our measures of nursing surveillance, the role of nursing in our overall quality of care measurement is not limited to those items found in our nursing surveillance scales. Nursing surveillance provided key data for our patient sickness at admission measure, our sickness at discharge measure, and our measures of in-hospital complications. Nursing surveillance data were also heavily used in the "if" parts of our process measures. If we stated, for example, that *if* the patient had hypotension, or confusion, *then* the patient should be in an intensive care unit, we used both physician and nursing notes to establish the *if*. Overall, because nursing notes are more voluminous than physician notes, we allocated about half of the time required for medical record review (approximately 75 minutes per record) to the review of nursing notes.

To collect data for the study, we trained experienced nurse abstractors over an intensive 14 day training session. Nurses then went to study hospitals and reviewed the selected medical records. Data collection was reviewed by a nurse supervisor (20 minutes per record on average), and portions were rereviewed by a physician (5 minutes per record on average).

We assessed interrater reliabilities for each disease using both records known to the data collectors to be test cases ("known", or "gold standard" records) and records on

which they did not know they were being monitored ("unknown", or "field data" records). There were 10 "known" records that were rated by 47 different data collectors. There were 162 different "unknown" records rated by two different data collectors.

Data Analysis

Scoring process criteria: We developed scores for process criteria based on clinical literature and judgement. Scores were calculated based on the percent of applicable criteria (points) accumulated. Scores were normalized into percentages of the standard deviations for items and scales (Kahn KL, Rogers WH, et al. 1990).

Developing Process Scales: To develop process scales, we first created clinically homogeneous groupings of items. We then evaluated our groupings by testing interitem reliabilities (Cronbach's alpha) for each scale, by comparing item placement with that suggested by a Likert (multitrait) scaling analysis, and by evaluating the correlation coefficients between scales. Final scales were: MD Assessment[1]; RN Assessment[2] (nurse surveillance); Technical Diagnostic; Technical Therapeutic, and Monitoring with ICU or Telemetry. An Overall Process Scale was created by combing scores on the five scales just mentioned. The MD Assessment and RN Assessment scales are further composed of Initial Assessment (assessing hospital care on days 1 and 2) and Subsequent Assessment (assessing care on day 3 or later of the hospital stay) components.

Measuring Sickness at Admission

We used clinical judgement to identify those patient characteristics present at admission that were likely to influence patient outcomes. We validated our clinical judgments by regressing death on sickness at admission.

Measuring Mortality

To determine whether patients had died after their hospitalizations, we matched our patients to social security files. We were able to match 92% of our sample. We measured death rates at 30 and 180 days after hospital admission, whether or not the patient was still hospitalized or had returned to the community.

RESULTS

Sickness at Admission

Our final Sickness at Admission scales for each of our five diseases were successful in predicting death at 30 and 180 days after hospital admission for our five diseases (Keeler, Kahn et al., 1990).

Process Criteria

Our individual nursing surveillance (RN assessment) criteria, and the rates of adherence to them, are presented in Table 1. Criteria consisted of selected activities that a nurse performs to monitor the clinical status of a patient. They included both physical examinations (e.g., measuring blood pressure at least three times per day on day 2 of the hospital stay in acute myocardial infarction patients) and gathering historical data (e.g., noting the presence or absence of shortness of breath in patients with pneumonia or congestive heart failure). Overall adherence to our criteria ranged from 48% (pupils assessed day 2 in CVA patients) to 98% (preadmission residence noted for hip fracture patients). Interestingly, although the nursing literature emphasizes functional status as an important nursing consideration, particularly in the elderly, nurses did not consistently assess functional status at admission. Urinary continence status was noted in only 64 - 71% of patients, and ambulatory status was noted in only 73 - 88%.

Reliability and Validity of Process Measures

As viewed by our panels, the process measures had face validity in terms of the importance of the individual criteria and their ability to sample important areas of in-hospital process. Interrater reliabilities were all above the .4 level that we considered acceptable for group

comparisons. Across diseases, item level Kappa scores for "known" records averaged .80. "Unknown" records averaged .78. Interrater reliabilities for the sample of nursing items that were evaluated are listed in Table 2. We tested scale reliabilities using Cronbach's alpha test. Reliabilities for the nursing surveillance scales (RN Assessment) range from .65 (acute myocardial infarction) to .36 (pneumonia), as listed in Table 3.

We evaluated construct validity by testing the correlations between scales. We would expect some correlation between process scales, because, for example, good physicians and good nurses may be more likely to practice in the same hospitals, and physicians who assess patients well may be more likely to treat them well. However, we would also expect that the correlations between physician and nurse scales would be only modest, because physician and nursing practices are substantially independent. The correlation between initial nursing surveillance (RN assessment) and subsequent nursing surveillance, on the other hand, should be high. As shown in Table 4, our expectations were met. There is only a .16 correlation between Initial MD Assessment and Initial RN Assessment (expected would be .45, based on reliabilities, if the two scales were measuring the same concept). On the other hand, there is a correlation between Initial Nursing Surveillance and Subsequent Nursing Surveillance of .30 (expected would be .24), indicating that these scales do measure similar concepts. Interestingly, there is a very low intercorrelation between both Initial MD Assessment and RN Initial Assessment Scales and the amount of time it took for patients to receive their first antibiotic doses. Though from a clinical point of view we might expect these to correlate, they do not, perhaps because they reflect other aspects of hospital process. Items that do not correlate well with a scale are best studied separately from it, and we have constructed our scales accordingly.

Process Outcome Link

We evaluated the predictive validity of our measures by determining whether patients who received worse care by our measures had worse outcomes. The quality of nursing surveillance was strongly linked to patient outcomes, controlling for sickness at admission, as shown in Table 5. The 30 day post admission mortality rate among patients with congestive heart failure with good nursing surveillance was 11% versus 17 percent for patients with poor surveillance. Likewise, acute myocardial infarction patients with good nursing surveillance had a 24% mortality while patients with poor surveillance had a 27% mortality rate. Among pneumonia patients mortality was 15% for good versus 19% for poor surveillance; among patients with cerebrovascular accident, mortality was 19% versus 24% for poor surveillance; and among patients with hip fracture, mortality was 4% versus 6% for poor surveillance. All of these relationships were significant at the $p<.05$ level except for the hip fracture results, which were not significant.

Nursing Quality Before and After PPS

For all diseases, the quality of nursing surveillance improved after the implementation of PPS. The changes in nursing surveillance for the four medical diseases were significant at the $p<.05$ level and translate into a potential mortality reduction of between .6 and .8% at 30 and 180 days between the before and after periods (Table 6). The surgical disease (hip fracture) had a relatively low mortality (4%) and the change in mortality potentially attributable to change in nursing surveillance for this disease was only about .3%, despite a significant improvement in nursing surveillance. Because the percentage reductions in mortality do not account for correlations between scales, they cannot be compared across scales; for example, the importance of the technical-diagnostic scale cannot be compared to that of the physician assessment scales or the nursing surveillance scales.

Overall, the quality of nursing surveillance behaved similarly to physician quality measures in terms of the change between the before PPS and the after PPS periods. Mortality rates, adjusted for sickness at admission, were lower during the post-PPS period when both physician assessment and nursing assessment were performed better, an improvement from the pre-PPS period when higher mortality rates were associated with lower (worse) scores for physician and nurse assessment. Overall, improvements in the quality of care across the five medical conditions were associated with a 1% reduction in 30 day mortality rates between 1981-1982 and 1985-1986.

DISCUSSION

The institution of PPS was accompanied by dramatic decreases in length of stay (Guterman et al., 1988) These decreases indicated the extent to which the new incentives encouraged hospitals to economize based on fixed reimbursements per case treated. Similar motives might encourage hospitals to economize on the intensity of nursing care, thus reducing the cost per case. Because nurses represent the greatest portion of cost per hospital day, this scenario seems especially likely. Measurement of nursing quality was therefore a critical component of our quality of care study.

How then do our nursing quality criteria fit into the overall conceptual framework for nursing processes? The nursing literature describes nursing process as consisting of assessment, nursing diagnosis, care-planning, intervention, and evaluation. Nursing interventions may be of two kinds: independent functions, and collaborative functions with other disciplines, primarily with physicians. When performing a collaborative function such as monitoring a patient's status, the nurse expects that her or his assessment will provide data for another type of professional to act upon more definitively. The fact that the function is collaborative does not mean that the assessment is not carried out independently by the nurse, nor that she or he is not professionally responsible for performing that function. Our nursing surveillance criteria reflect some of the many actions that the nurse performs to monitor the clinical status of the patient; the nurse is held accountable for performing these actions whether or not there is an order for them.

Nursing criteria used in this study focus on surveillance activities, because these tend to be measurable in the record, and because they might be expected to be strongly linked to mortality as an outcome. Components of nursing diagnoses and defining characteristics (Brunner, 1984; Chang, 1988; Gordon, 1982; Kim, 1989; McFarland and McLane, 1987) we focussed on included functional status at admission and discharge, nutrition (CVA), and alteration or potential alteration in respiratory function (CHF, PNE), fluid status (AMI), cardiac rhythm (AMI), chest pain (AMI), blood pressure (CHF, PNE, CVA), and neurologic status (CVA). We hoped that the items we measured in each of these areas sampled the nurse's behavior accurately enough to reflect his or her degree of attentiveness to the patient's clinical situation. Others have defined a broader range of nursing quality measures (Hegyvary and Hussman, 1976; Goldman, 1990; Long and Clinton, 1984; Pearson, 1989) which have not been addressed in the present study.

We found evidence that our measures did reflect overall nursing attentiveness. When nurses did poorly by our measures, their patients were more likely, at all levels of sickness at admission, to die. This link between the quality of nursing surveillance and the outcomes of care, which occurred for all study diseases except hip fracture demonstrates the critical role played by nurses in determining the patient's clinical outcomes. In addition, nursing assessment is probably linked to other critical aspects of the nurse's role. For example, nurses who monitor shortness of breath to assess the clinical condition of a patient with congestive heart failure may be also be attentive to other aspects of the patient's condition,

whereas nurses who fail to perform the critical initial step of assessing the patient may be likely to fail in other aspects of nursing process as well.

We found that nursing surveillance scores improved between 1981-1982 and 1985-1986. This finding paralleled our findings regarding the other types of quality of care we measured. Improved medical care process over this time period was associated with a decrease in the death rate, despite the increased sickness at admission of patients in 1985-1986 (Keeler E, Kahn KL, et. al., 1990).

A limitation of our study was that, although our study team included nurses, our criteria development panels were carried out with the help of physicians in Peer Review Organizations, and we were not funded to enlist nursing panel members. Another study limitation is that the patient outcome we most emphasized in developing and evaluating our criteria was death, either in the hospital or soon following the hospitalization. This was because our study was not designed to focus on the other major type of health status outcome—post hospital functional status or quality of life. The rehabilitative aspects of nursing comprise another dimension of nursing process (Kitson, 1986). A study to focus on functional status outcomes, however, would have been prospective, and would not have fulfilled the requirements of our proposal. Related to this limitation is another—the requirement that our process measures be reliably attainable through medical record review. In many cases, we discarded potentially important measures of nursing quality because of concerns that the criteria might be better documented in some hospitals than others or in the post rather than the pre PPS period, or simply were not reliably documented at all. We hypothesized that key elements of the nursing assessment, such as determination of functional status, and assessment of key signs and symptoms of disease, would be most often recorded if performed, while actions such as counseling patients might be recorded or not, depending upon the situation and hospital.

In conclusion, this study has demonstrated the importance of measuring nursing surveillance functions in studies designed to determine the quality of care for major medical diseases. In addition, we have found that, despite the implementation of the PPS system, nursing surveillance has continued to improve over time. Future studies are needed to establish the relationships between measures of nursing surveillance functions and other aspects of nursing care; to determine the effects of nursing interventions on outcomes other than death; and to monitor nursing performance over succeeding years as financial constraints continue to place pressure on the process of medical care.

REFERENCES

Brunner LS, Suddarth DS. Textbook of Medical-Surgical Nursing, Philadelphia, PA: J. B. Lippincott. (1984).

Chang BE, Brazel-Villanueva DW, Iverson R, "Self-care deficit with etiologies: reliability of measurement", Nursing Diagnosis 1990; 1(1):31-36.

Chang BL, Gonzales E, Caswell D, "Validity and reliability of an assessment guide for identifying nursing diagnoses", The Australian Journal of Advanced Nursing 1988; 5(2): 16-22.

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," Journal of American Medical Association, 1990;264(15):1956-1961.

Fitzpatrick JJ, "Conceptual basis for the organization and advancement of nursing knowledge: nursing diagnosis/ taxonomy", Nursing Diagnosis 1990; 1(3): 102-106.

Goldman RC, "Nursing process components as a framework for monitoring and evaluation activities", Journal Nursing Quality Assurance 1990;4(4): 17-25.

Gordon M. Nursing Diagnosis, New York, NY: McGraw-Hill. (1987).

Halloran, EJ. Nursing workload, medical diagnosis-related groups, and nursing diagnoses. Research in Nursing and Health,8, 421. (1985).

Hegyvary ST, Haussmann RD. Correlates of the quality of nursing care. Journal Nursing Administration, 6, 22-27. (1976a).

Hegyvary ST, Haussmann RD. Relationship of nusing process and patient outcomes. Journal Nursing Administration, 6, 18-21. (1976a).

Kahn KL, Chassin MR, Rubenstein LV, et al., Medical Record Abstraction Form and Guidelines for Patients Assessing Quality of Care for Hospitalized Patients with Congestive Heart Failure, The RAND Corporation, N-2798-HCFA.

Kahn KL, Keeler EB, Sherwood MJ, et al., "Comparing Outcomes of Care Before and After Implementation of the DRG-Based Prospective Payment System," Journal of American Medical Association, 1990;264(15):1984-1988.

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," Journal of American Medical Association, 1990:264(15):1969-1973.

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," Journal of the American Medical Association, 1990;264(11):1953-1955.

Keeler EB, Kahn KL, Draper D, et al., "Changes in Sickness at Admission Following the Introduction of the Prospective Payment System," Journal of American Medical Association, 1990;264(15): 1962-1968.

Acute Myocardial Infarction, The RAND Corporation, N-2799-HCFA.

Lancour J, "Nursing diagnosis and the effectiveness initiative", Nursing Diagnosis 1991; 2(1): 33-35.

Lang NM and Clinton JF. Assessment of Quality of Nursing Care. HH Werley and JJ Fitzpatrick (eds.). Annual Review of Nursing Research. New York Springer; 1984;135-163.

McFarland GK, McFarlane EA. Nursing Diagnosis and Intervention Planning for Patient Care, St. Louis, MI, C.V. Mosby Company. (1989).

Pearson A, Durant I, Punton S, "Determining quality in a unit where nursing is the primary intervention", Journal Advanced Nursing 1989; 14: 269-273.

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," Journal of American Medical Association, 1990;264(15):1989-1994.

Roth CP, Kahn KL, Sherwood MJ, et al., Medical Record Abstraction Form and Guidelines for Patients Assessing Quality of Care for Hospitalized Patients with Pneumonia, The RAND Corporation, N-2801-HCFA.

Roth K, Distefano JJ, Chang BL, "Development of the automated nursing assessment tool", Computers in Nursing 1989;7(5):222-227.

Rubenstein LV, Kahn KL, Reinisch EJ, et al., "Changes in Quality of Care for Five Diseases Measured by Implicit Review 1981 to 1986," Journal of American Medical Association, 1990;264(15):1974-1979.

Rubenstein LV, Kosecoff J, Kahn KL, et al., Medical Record Abstraction Form and Guidelines for Patients with Cerebrovascular Accident, The RAND Corporation, N-2802-HCFA.

Sherwood MJ, Kahn KL, Kosecoff J, et al., Medical Record Abstraction Form and Guidelines for Patients Assessing Quality of Care for Hospitalized Patients with Hip Fracture, The RAND Corporation, N-2800-HCFA.