# PMC Patient Severity Scale: Derivation and Validation

*Wanda W. Young, Susan Kohler, and Jeanne Kowalski*

**Objective.** This study describes the derivation and validation of the Patient Management Category (PMC) Severity Scale, which provides a method of assessing the overall severity of a hospitalized patient's illnesses, based on the patient's unique clinical conditions, their interaction, and the resultant, combined risk of morbidity and mortality.

**Data Sources.** Derivation of the PMC Severity Scale was based on clinical judgment together with empirical analysis of more than a half million patients discharged from acute care hospitals in Maryland during 1989. The scale was validated by using two distinct calendar years (1988 and 1990) of patient data from the same Maryland hospitals and a six-month patient database from California (1990).

**Study Design.** The PMC Severity Scale is an ordinal scale with seven levels: Level 7 represents the greatest likelihood of death and major disease burden. The scale quantifies the severity of each of the patient's disease(s) and accounts for the effect of all coexisting conditions and complications.

**Data Extraction Methods.** Publicly available, statewide all-payer claims databases were acquired from Maryland and California.

**Principal Findings.** The independent relationships between the PMC Severity Scale with mortality and with length of stay are statistically different across severity levels within each population tested, but the relationships are statistically similar over time. Further, the PMC Severity Scale was determined to be a stable predictor of mortality and LOS across two diverse geographic regions.

**Conclusions.** Since the severity of a patient's illness is one of the factors that influences the outcomes of care, the PMC Severity Scale can be used successfully as a risk adjustment tool in a variety of quality applications.

**Keywords.** Severity of illness, risk adjustment, outcomes assessment, comorbidity, scale, validation

Consumers, employers, and regulators alike have increased their demands in the past few years for accurate comparisons of health care quality. Although their shared goal has been to ensure that appropriate, cost-effective

care is delivered consistently to all patients, there is less consensus on how to measure whether or not we are achieving that goal. The specific task of measuring the appropriateness and outcomes associated with certain health care services is complicated by the fact that each patient may be unique in his or her combination of diseases and risk factors at any given point in time. To measure these severity of illness distinctions, this study describes the development and validation of the Patient Management Category (PMC) Severity Scale, a severity measure to improve the precision of patient outcomes research and assessments of provider performance.

The PMC Severity Scale is an adjunct to the existing Patient Management Category (PMC) Classification and Intensity Scoring System, a computerized patient classification and scoring system that is both clinically specific and linked with different levels of care (Young, Kohler, and Macioce 1992). The PMC system was originally developed with a grant from the Health Care Financing Administration to the research staff of Blue Cross of Western Pennsylvania (Young, Joyce, Schuchert, et al. 1985). The original system was made available through the National Technical Information Service, while the maintenance and dissemination of the system has been managed by The Pittsburgh Research Institute, a 501(c)(3) not-for-profit health services research organization.[1] As part of this ongoing enhancement, we sought in this project to create an independent method for quantifying the overall severity of each patient's unique combination of diseases, interaction of the diseases with effective treatment, and the resultant combined risk of morbidity and mortality. The goal was to develop a severity index associated with PMCs that could be used directly to improve the standardization of outcome measures and to complement the PMC Relative Intensity Score (PMC-RIS), which is useful for adjusting costs.

Before describing the PMC Severity Scale, the PMC Classification and Intensity Scoring System as well as some general issues in designing patient scoring systems will be reviewed. These two areas will provide a framework for assessing potential applications of the severity measure described in this article.

---

Address correspondence and requests for reprints to Wanda W. Young, Sc.D., President, The Pittsburgh Research Institute, Fifth Avenue Place, 120 Fifth Avenue, Suite 1711, Pittsburgh, PA 15222-3099. Susan Kohler, R.N., B.S.N., is Manager, Clinical Studies, The Pittsburgh Research Institute; Jeanne Kowalski is Research Analyst at The Pittsburgh Research Institute. This article, submitted to *Health Services Research* on May 24, 1993, was revised and accepted for publication on November 5, 1993.

## BACKGROUND

The original objectives of developing the Patient Management Categories and the PMC Relative Intensity Scoring System were twofold: (1) to identify, in a clinically specific way, the types of patients treated at different institutions, and (2) to measure severity and resource intensity differences among those patient types. The goal was to design a system that is both clinically meaningful and useful in predicting expected hospital resource use and costs. To accomplish these objectives, the following analytic tools were developed:

**Patient Management Categories.** A *computerized, diagnosis-based patient classification,* developed with extensive clinical input from physician panels, that incorporates severity of illness distinctions and defines comorbidity;

**Relative Intensity Scores.** A set of cost-based relative weights that (1) reflect the intensity of hospital services required to manage a clinically specific patient type in relation to the average hospitalized patient, and (2) are combined in a computerized methodology to yield *one Relative Intensity Score (RIS) for each patient* based on that patient's particular disease conditions (single disease or comorbid) and complications; and

**Patient Management PATHs.** Computerized, *physician-specified clinical management strategies* (one for each PMC), each of which consists of diagnostic and treatment services for effective care of the typical patient in that PMC.

## PATIENT MANAGEMENT CATEGORIES

PMC Classification (*Release 5.0*; see Young, Kohler, and Macioce 1992) consists of 830 Patient Management Categories that describe all patients treated in general acute care hospitals. In addition to the categorization of general medical and surgical diseases (e.g., diabetes, pneumonia, vascular emboli, AIDS), special populations (e.g., neonates and deliveries), and psychiatric disorders, patient categories have been defined for specific complications as well (e.g., septicemia, wound dehiscence, hemorrhage).

PMCs were originally defined, by physicians, in clinical terms independent of historic patient data. Extensive physician consultation was obtained through the formation of more than 50 disease-specific panels, each of which consisted of four to six physicians (both generalists and specialists) who treated patients with the disease(s) being modeled in that panel session

(Young, Joyce, Schuchert, et al. 1985). Initially, more than 125 physicians from southwestern Pennsylvania participated in these expert panels. Since its original development, however, many more physicians from the United States and abroad have contributed to the research that has been part of the ongoing modification and enhancement of the PMCs and PATHs.

After patient types were designated by physician panels, codes in the *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) were mapped to the categories to computerize the classification. The resultant PMC Classification software uses the unique combination of ICD-9-CM diagnosis and procedure codes recorded on a patient's computerized discharge abstract to assign one or more clinically specific PMCs to that patient.[2] Unlike other classifications that are driven by the principal diagnosis code listed on the patient's abstract (such as DRGs), PMCs disregard the sequence or order in which the diagnosis codes are listed to identify all comorbid conditions as well as specific complications. It is the way in which these codes are aggregated and the interrelationship of diagnosis codes that is critical to the accurate identification of clinical patient types and the valid assignment of PMCs to patients.

An example will illustrate the importance of recognizing the relationship among diagnosis codes for accurate classification. Table 1 shows the comorbid conditions (two PMCs) of one patient with five ICD-9-CM diagnosis codes. Two codes describe the clinical manifestation (perforation) of the patient's diverticular disease while three of the codes listed are related to the patient's AMI and related complication (cardiogenic shock). The Patient Management Category computerized algorithm searches the list of diagnosis codes to determine the general disease(s) that were treated in that hospitalization–in this case, Diverticular Disease and Acute Myocardial Infarction (AMI). Within each of these disease areas, which are modules or subroutines of the PMC Software, the combination of related diagnosis codes is recognized by the software and used to make the clinically specific PMC assignment(s). The same two PMC assignments would be made for this patient regardless of the order of the diagnosis codes recorded on the patient's computerized record.

The term "comorbid" means that more than one *disease* or *pathological condition* is present, not that there is more than one *diagnosis code* present. It should also be emphasized that comorbid patients are not necessarily more severely ill or more costly to manage than patients with a single disease. That is, a patient can be comorbid with two relatively minor conditions, each requiring few hospital resources for effective management. Therefore, the identity of the particular comorbid conditions will determine the severity of the patient's illness, the intensity of resources required to managed that patient, and the resultant cost of care.

Table 1: Example of PMC Assignment Process

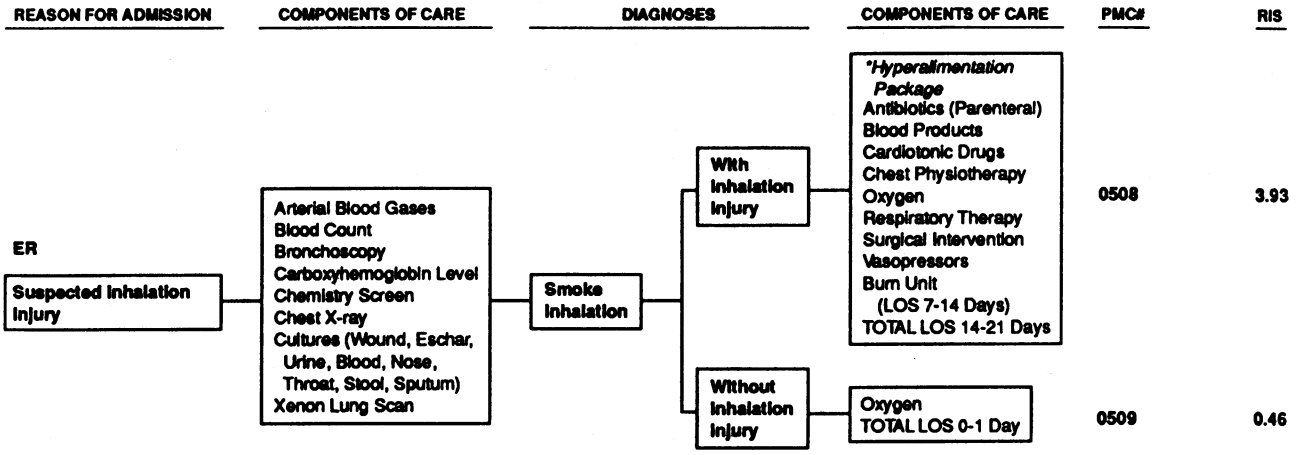
| *ICD-9-CM Diagnosis* | *PMC Assignment* | *PMC-RIS* |
|---|---|---|
| 569.83 Perforation of intestine<br>562.11 Diverticulitis of colon | 0104 Diverticular Disease: Peritonitis/Perforation | 1.79 |
| 427.81 Sinoatrial node dysfunction<br>410.11 Acute myocardial infarction of other anterior wall<br>785.51 Cardiogenic shock | 0308 AMI: Cardiogenic Shock | 2.76 |
| Comorbid combination | | 3.75 |

Physicians also identified a subset of PMCs that are conditions likely to complicate the management of the patient and influence patient outcomes, both morbidity and mortality. These PMCs are referred to as complication PMCs even though they include a broad range of conditions that complicate medical management and are not limited to complications that result from treatment of the patient. It should be noted that, with most administrative data (and even with medical record review in some instances), it is extremely difficult to determine when the complication occurred (i.e., after admission or concomitant with the condition that led to hospitalization) and what precipitated the complication (e.g., the patient's compromised status or the treatment provided). Nevertheless, it is possible with administrative data to identify these complicating factors and to use the designated PMCs to assess the impact of complications on patient morbidity and mortality as we have done in this study.

## PATIENT MANAGEMENT PATHS

During the original development, panels of physicians not only identified PMCs within each disease area, but they also specified the hospital services required for the effective treatment of a typical patient in each PMC. Each of these physician-specified management strategies, referred to as Patient Management PATHs, consists of diagnostic and treatment services (e.g., specific x-rays, scans, laboratory studies, and operative procedures required, if any) as well as expected lengths of stay in special care units and in total. These PATHs represent effective and efficient resource use for a typical patient in each category; they do not represent a prescription for care of a particular patient. An example of the form that this takes is shown in Figure 1.

Figure 1: Sample Patient Management PATH–Burn: Smoke Inhalation



The PMCs shown in this example are PMC 0508, Burn: Smoke Inhalation with Inhalation Injury, and PMC 0509, Burn: Smoke Inhalation without Inhalation Injury. Both smoke inhalation patient types require the availability of an emergency room and are typically managed with oxygen, and with laboratory and radiology studies. Additional resources, however, are required for PMC 0508, Smoke Inhalation with Inhalation Injury, including drug therapy, respiratory therapy, possible surgical intervention, and the availability of a special care unit for 7 to 14 days. The total length of stay (LOS) expected for this inhalation injury patient is 14 to 21 days in contrast to only 0 to 1 day of acute care for the smoke inhalation patient without inhalation injury. Physicians have, in fact, designated the latter PMC as a potential ambulatory patient type.

PMCs and PATHs are both conceptually and operationally distinct. The PMCs in this example are defined based on combinations of diagnosis codes (in any sequence) recorded on patient discharge abstract data. By contrast, the services on each Patient Management PATH provided, for each patient type, the basis for the identification of hospital costs, which were then used to derive a cost-based relative value scale for PMCs, reflecting the relative intensity of expected resource requirements for each patient type. Specifically, patient-related hospital costs (derived through detailed cost finding) were identified for each component of care and then accumulated for each Patient Management Category to determine the expected cost of managing that patient type.

## PMC RELATIVE INTENSITY SCORE

The costs of the services specified on each PATH (one for each PMC) were translated into cost-based relative weights, and are referred to as PMC Relative Intensity Scores (PMC-RIS). The significant difference in resources required in the management of these two burns is reflected in their category weights shown at the far right on Figure 1. That is, managing a patient with an inhalation injury is more than eight times as costly as managing a smoke inhalation patient without inhalation injury.

The example shown indicates the method used to assign a PMC-RIS to a patient with a single PMC assignment. Recall, however, that the PMC computerized algorithm permits a single patient to receive multiple PMC assignments to reflect multiple injuries, specific comorbid conditions, and complications. When this occurs, one PMC Relative Intensity Score for the patient is derived by merging the cost-based weights associated with the overlapping services (counting them only once) and combining them with the cost-based weights of the unique services associated with each condition.

(As part of the PMC Classification System, software exists to carry out this scoring process automatically.) This methodology yields one PMC-RIS for the patient that is not additive of the individual PMC scores, but instead represents an adjustment upward to reflect the increased resource use expected for patients with certain comorbid conditions, complications, or both.

Table 1 shows the upward adjustment for a patient who had diverticular disease with peritonitis and an acute myocardial infarction (AMI) with cardiogenic shock. If this patient had only the peritonitis, the PMC-RIS would have been 1.79, reflecting the surgical management required for that condition alone. A patient with an AMI and cardiogenic shock, without other conditions, would be assigned a PMC-RIS of 2.76, reflecting more resource-intense management. A patient with this particular combination of diseases, however, is assigned a PMC-RIS of 3.75, reflecting the increased resource use that is expected for this particular combination of clinical conditions.

This classification and scoring process has been applied to hospitalized patient records in various hospital and regional databases.[3] For 99 percent of all these patients, PMC Relative Intensity Scores range from 0.05 (reflecting ambulatory management) to 4.4 (reflecting a more severe patient type requiring more intense and costly hospital resources to manage). Approximately 95 percent of patients generally have a PMC-RIS below 2.6. A score of 1.0 reflects the average cost patient in the population base. Since the amount of the adjustment associated with comorbidity and complications depends on the unique combination of conditions that exist for a patient, there is no upper limit on the weighting scale. In practice, the maximum value in all of the databases analyzed to date is 17.2.

In this way, PMCs define clinically specific case types and can be used to differentiate patients with different resource needs (as opposed to differentiating patients based on resource use). Severity distinctions were made by physicians in defining categories, and thus, severity is implicitly measured in the PMC classification system. Because PMC Relative Intensity Scores are based on the relative costs of treating patients of different severity and with different resource needs, and because they incorporate the effect of comorbid conditions and complications, they are an interval scale that can be used across diseases to adjust and predict costs. They have also been used as a surrogate measure for severity in morbidity and mortality analyses.

## ISSUES IN THE DERIVATION OF RELATIVE SCORING SYSTEMS

The categories in a diagnosis-based classification represent a nominal form of measurement. To use such a classification effectively in comparative

analyses across different populations and to adjust these populations for expected differences among patients with respect to their illnesses, it is often necessary to aggregate these patient categories into some form of composite measure. Such a measure can be based on one or more of a number of dimensions: severity, intensity of resource use or need, duration of days required for treatment, and actual or expected costs are a few of them. Depending on the basis of the index, resources (e.g., days, services, or costs) required to manage diverse patient populations can be projected and outcomes (e.g., morbidity or mortality) can be estimated.

Given the variety of measures possible, the decision to use a particular relative value scale should be guided by its intended use (The Hospital Research and Educational Trust 1989). Common practice among health care researchers, however, has been to use severity measures to adjust hospital expenses, assuming that "severity of illness," a concept that is generally not very well defined, is directly related to what it "should cost" for effective care. That is, researchers and practitioners have assumed that the more severely ill a patient is, the higher the costs should be.

Although severity and intensity of resource use and costs are directly related in many instances, in some disease groups this is clearly not the case. For example, an AMI patient with cardiogenic shock is one of the most severe PMCs (with an expected in-hospital death rate greater than 80 percent), but not one of the most costly cardiac patient types. Not only do patients with this condition frequently die early in their hospitalization, but those who are discharged alive have typically been managed medically (with or without angioplasty as opposed to open heart surgery). Other AMI patients with a lower in-hospital death rate (e.g., AMI with congestive heart failure) are more costly because they are more likely to receive major operative procedures. Thus, it should not be assumed for a particular disease that severity per se will necessarily be directly related to costs, charges, length of stay, or any other measure of resources used in patient management.

Despite the development of numerous indexes that purport to measure severity of illness, much ambiguity remains about the assumptions underlying the development of these measures (Stein, Gortmaker, Perrin, et al. 1987), as well as their construct and predictive validity. Not even the definition of severity is as straightforward as one would think given the commonly held belief that severity of illness is an important factor in determining patient morbidity and mortality (Thomas and Longo 1990).

Physicians seem to agree intuitively that the more severe clinical conditions are those associated with a greater probability of immediate death or disability and/or whose management (diagnosis and treatment) is more complicated than that for other conditions.[4] As part of the development of PMCs, physicians recognized severity distinctions among clinically distinct

patient types within a disease and were able to rank these patient types by degree of severity. For example, a diverticular disease inpatient with nonmassive gastrointestinal bleeding was considered less severe than a diverticular disease patient hospitalized because of an obstruction or fistula. These types of rank-order judgments within a disease have been incorporated into the modular hierarchy of the computerized PMC classification (Young 1984) and in other disease classification systems as well (Gonnella, Hornbrook, and Louis 1984).

The difficulty arises when one tries to quantify these severity differences, especially across different diseases. An additional complexity arises when two or more clinical conditions occur simultaneously. That is, does the additional disease and/or complication make the patient more severe? To what extent is the probability of continued morbidity or imminent death increased by the presence of more than one disease or clinical manifestation of the disease(s)?

In addition to these disease-specific issues are other aspects of a severity index that should be identified before it is selected for use. For example, is the severity index attempting to predict the probability of death, the potential for organ failure, the risk of permanent impairment, the total impact of the disease process on the patient's long-term survival probability, or some combination of morbidity and mortality? Does the index refer to disease severity or patient severity at a point in time? What is the methodology (i.e., statistical analyses versus psychometric methods) used to derive final severity scale values? Similarly, it is important to define the reference disease(s) or clinical condition(s), the applicable time interval, and the extent to which treatment and its interaction with the disease are included in the conceptualization of severity.

A severity index designed to standardize hospital lengths of stay and death rates should incorporate the patient's acute illnesses during the hospital stay as well as chronic or coexisting conditions that have the potential to influence the patient's overall probability of continued morbidity or death (Charlson et al. 1987; Dubois et al. 1987; Greenfield et al. 1988). Complications that are not preventable and/or are part of the disease process should also be included in the construction of such a relative severity score. In most cases, both the severity of a patient's illness and the types, quantity, and intensity of effective care rendered to the patient are factors that influence the outcomes of care, especially the probability of long-term survival.

This article reports on the development of the PMC Severity Scale, a seven-level ordinal scale that quantifies the severity of a patient's clinical condition(s) and/or the patient's clinical manifestation(s) of disease (not necessarily the disease per se) during the hospital stay. This clinical severity level is then adjusted upward for the effect of specific comorbid conditions

and complications, yielding one overall patient severity level (PMC Severity Score) for that hospitalization.

## DATA AND METHODS

The methodology used to develop the PMC Patient Severity Scale was based primarily on an empirical analysis of more than a half million patients discharged from acute care hospitals in Maryland during calendar year 1989. Patient discharges from the same Maryland hospitals during calendar years 1988 and 1990 were used to assess the reliability of the PMC Severity Scale over time. A six-month patient database from California (1990) was also used to determine the validity of the PMC Severity Scale across geographic regions.

A two-part strategy was designed to focus first on clinical and empirical distinctions among *diseases* (PMCs), and then on an assessment of the *impact of multiple diseases on a particular patient's hospitalization.* The first stage of analysis resulted in the assignment of each PMC to one of four levels of severity. In the second stage, this scale was expanded to a seven-point scale and redefined to incorporate the effects of comorbidity and complications. Thus, the final PMC Severity Scale measures the severity of the patient's illness episode rather than the severity of each individual disease.

## RESULTS

### *Assignment of Severity Levels to PMCs*

The first step in the process of deriving a relative numerical value to measure the severity of each patient's hospitalization was to focus on *diseases* as if they were managed singly, that is, without regard to comorbidity or other patient characteristics. Thus, a rank of 1 to 4 was assigned to each of the 830 PMCs, based on the subjective clinical judgment of the nurse researcher/author. Level 1 represented the lowest expected severity, that is, those patients who could potentially be treated in an ambulatory setting as well as those with other minor medical problems. Level 4 represented the highest expected severity and life-threatening situations. This was a way to derive a preliminary scale that could be analyzed empirically.

Since severity is not directly quantifiable, surrogate outcomes (death rates and LOS) were used to assess and adjust this initial scale. Most investigations of severity classify disease categories by average death rate to reflect the probability of death and risk of organ failure. Since relatively few

diseases typically result in death, however, in this analysis the LOS of each disease-specific hospital stay was also chosen as a surrogate for morbidity.

Average lengths of stay for single diseases (PMCs) were derived from the all-payer 1989 Maryland statewide database of approximately 600,000 patient discharge records. Patients who were transferred at discharge to other short-term hospitals, were discharged against medical advice, or had a hospital stay greater than 105 days were excluded from these calculations, leaving 568,762 cases in the analysis.

Recall that a Patient Management PATH, associated with each PMC, provides physician-specified information about the nature of the types of admissions associated with the particular disease (emergent, urgent, or elective) as well as physician-specified expectations regarding the need for diagnostic and therapeutic interventions and continuous monitoring. These clinical expectations, combined with data on the probability of hospital death and LOS associated with each PMC, were used to adjust the initial severity level assignments made for each PMC and to ensure that these assignments were made consistently across diseases and body systems. Although the resultant disease-specific severity scale (from 1 to 4) explained approximately 28 percent of the variation in LOS for single-disease patients in the developmental database, this was only an intermediate step toward deriving an overall score reflecting the severity of all of the clinical conditions managed in a patient's hospitalization.

*Assignment of Severity Levels to Patients*

The four disease-specific severity levels (defined earlier in the preliminary scale) might be adequate to use in predicting and/or adjusting LOS for those patients who have only one disease that is managed during a hospitalization. In the Maryland statewide database, however, as in other large populations, approximately 43 percent of the patient records received two or more PMC assignments, indicating that more than one morbid condition, sometimes together with additional complications, were managed in a single hospitalization. In such situations, the problem is to define the overall severity of the patient, not just the severity of each clinical condition or disease.

The complex methodological question here is whether the combined severity levels that characterize each unique disease and/or complication of the patient has an additive, hierarchical, or interactive effect on the patient's morbidity (as measured by LOS) and outcomes (as measured by short-term survival). For example, for a comorbid case with PMC 303, AMI: Tachyrhythmia (severity level 3) and PMC 704, Renal: Acute Renal Failure without Dialysis (severity level 4), should the overall patient severity level be "7" (an additive effect)? or "4" (a hierarchical effect, with no additional

effect beyond the highest severity level)? or do the two diseases interact to produce an adjusted severity level that is higher than each treated separately but not as high as the additive effect? That is, is the combined effect (3 and 4) equivalent to the effect that a level "6" has on LOS (an interactive effect)?

To examine the impact of multiple diseases and/or complications managed in a single hospitalization, one or more PMCs were assigned to each patient record in the Maryland statewide database along with the severity level associated with each of the assigned PMC(s). Given that each patient record in this database could have received up to six PMC assignments, the number of possible severity level combinations is high, especially if multiple PMCs of the same severity level are considered separately. To make this a manageable analysis task, the combinatorial formula was first used to identify patterns of disease severity, assuming no duplication of severity level for a given patient. Next, among statistically distinct combinations, the remaining possible combinations were examined, this time analyzing multiple PMCs of the same severity level. By aggregating all possible combinations of severity levels and testing their distinctiveness with respect to LOS, an expanded severity scale incorporating the effect of comorbidity was developed.

### *Analysis of Severity Level Combinations, without Duplication*

Initially, therefore, to limit the number of subgroups in the first phase of testing, only unique severity level combinations were examined. Using the combinatorial formula to create all possible mathematical combinations of the four distinct severity levels resulted in 15 severity level combinations.[5] Each patient in the statewide Maryland database was assigned to one of these combinations based on the unique severity levels associated with that patient's PMCs. For example, patients assigned only the single PMC 5012, Urinary Tract Infection (severity level 1), were grouped into combination {1}. A patient assigned to multiple PMCs such as PMC 5012, Urinary Tract Infection (severity level 1); PMC 1101, Uncomplicated Appendicitis (severity level 2); and PMC 707, Subarachnoid Hemorrhage without Operation (severity level 4), were grouped into combination {1&2&4}.

Because, by definition, these mathematical combinations contain only distinct numerical values, comorbid patients with two or more PMCs of the same severity level were grouped, for this initial analysis, without regard to the frequency with which that severity level occurred. For example, a comorbid case assigned two clinically distinct PMCs, each with a severity level 2, was considered along with patients who had only one PMC with a severity level 2 (combination {2}). Similarly, a patient assigned three PMCs

reflecting severity levels 1, 1, and 3 was treated in this first phase of analysis as if that patient had only one level 1 and one level 3 (combination {1&3}).

To test the homogeneity of these severity level combinations with respect to LOS, distributions of LOS were analyzed for patients (discharged alive) who were assigned to each combination. In general, the LOS distributions were positively skewed for all combinations; lower severity level combinations, however, appeared to be more positively skewed than the higher severity level combinations. Therefore, the hypothesis of identical LOS distributions between and within severity level combinations was statistically tested using the Kruskal Wallis $H$-test (chi-square approximation), resulting from a nonparametric one-way analysis of variance. This hypothesis was rejected indicating that the data provided sufficient evidence to indicate that at least 2 of the 15 distributions differed in location. Using Scheffe's test for all possible pairwise comparisons, the combinations were further tested to determine which of the 15 were significantly different from each other.

Results of this analysis indicated that the four single severity level combinations (${}_4C_1$, in note 5) and the six combinations containing two unique severity levels (${}_4C_2$, in note 5) had a significantly different impact on LOS ($p < .0001$), both between combinations (e.g., severity level combination {2&4} was significantly different from combination {2} or {4} alone) and within combinations (e.g., severity level combination {1&2} was different from {1&3}). In general, the resultant distributions were positively skewed and exhibited different central tendencies and variations between and within combinations.

Combinations that were not statistically distinct generally included three or more severity levels. The LOS distributions for these combinations were not significantly different from the lengths of stay for patients assigned only the two highest severity levels in that combination. That is, patients assigned to severity levels 1, 2, and 4 had a LOS distribution similar to cases assigned to severity levels 2 and 4. Likewise, combinations {1&2&3&4}, {1&3&4}, and {2&3&4} were similar to patients assigned only to severity levels 3 and 4. As a result of this finding, comorbid patients whose conditions were characterized by three or more unique severity levels were reassigned to the combination defined by the two highest severity levels for that patient.

### *Analysis of Severity Level Combinations, with Duplication*

Next, all combinations of patients receiving two or more PMC assignments of the same severity level, within each of the ten statistically distinct combinations from the previous phase of testing, were analyzed. Specifically, using the multiplicative rule within these ten combinations and allowing duplication of the same severity level, 66 severity level combinations were

identified. Duplication of the same severity level was limited to three;[6] that is, if a patient had four PMCs, each with the severity level 2, the patient is represented in the combination {2&2&2}. To test for homogeneity with respect to LOS, distributions of LOS were examined and the same statistical analyses as in the previous phase of testing (Kruskal Wallis *H*-test and Scheffe's test) were conducted between and within all combinations of various sizes.

Based on this analysis, all combinations were reordered into seven severity levels. Table 2 illustrates the final composition of each of these seven severity levels. The result is a PMC Severity Scale which is an ordinal scale, with Level 7 representing the greatest likelihood of death and major disease burden. The scale quantifies the severity of each of the patient's disease(s) and accounts for the impact of all coexisting conditions and complications that are typically treated in general acute care hospitals.

## VALIDATION OF PMC SEVERITY SCALE

Although the development of the PMC Severity Scale incorporated both clinical judgment and statistical analyses, its overall content validity is primarily derived from the clinical framework of the Patient Management Category (PMC) Classification System. PMCs define patients' clinical conditions accurately relative to other diagnosis-based classifications (Thomas, Holloway, and Guire 1993; Young, Macioce, and Young 1990), identify specific

Table 2: Severity Level Combinations within the Final Seven Severity Levels (L)

| | | | | | | |
|---|---|---|---|---|---|---|
| Level 1 | 1L1; 2L1; 3L1; 1L2; 1L1 & 1L2; 2L1 & 1L2; 3L1 & 1L2 | | | | | |
| Level 2 | 1L1 & 2L2; 2L1 & 2L2; 3L1 & 2L2; 2L2 | | | | | |
| Level 3 | 1L1 & 3L2; 2L1 & 3L2; 1L3; 3L2; 1L1 & 1L3; 2L1 & 1L3; 3L1 & 1L3; 1L2 & 1L3 | | | | | |
| Level 4 | 2L1 & 1L4; 3L1 & 1L4; 1L1 & 1L4; 1L4; 2L3; 1L1 & 2L3; 2L1 & 2L3; 3L1 & 2L3; 2L2 & 1L3; 3L2 & 1L3 | | | | | |
| Level 5 | 1L2 & 1L4; 2L2 & 1L4; 1L2 & 2L3; 2L2 & 2L3; 3L2 & 1L4; 3L2 & 2L3 | | | | | |
| Level 6 | 1L1 & 2L4; 2L1 & 2L4; 3L1 & 2L4; 1L2 & 2L4; 2L4; 3L3; 1L1 & 3L3; 2L1 & 3L3; 3L1 & 3L3; 1L2 & 3L3; 2L2 & 3L3; 2L2 & 2L4; 1L3 & 1L4; 2L3 & 1L4 | | | | | |
| Level 7 | 1L1 & 3L4; 2L1 & 3L4; 3L1 & 3L4; 1L2 & 3L4; 2L2 & 3L4; 3L2 & 3L4; 1L3 & 3L4; 2L3 & 3L4; 3L3 & 3L4; 3L3 & 1L4; 1L3 & 2L4; 2L3 & 2L4; 3L3 & 2L4; 3L2 & 2L4; 3L2 & 3L3 | | | | | |

comorbid conditions and complications of each patient (Young 1984), and can uncover the clinical heterogeneity within other statistically homogenous patient groups (Charbonneau, Ostrowski, Poehner, et al. 1988).

Because the distinct concepts of clinical specificity, comorbidity, intensity, and severity of illness are separately operationalized in the PMC system, it has been possible to assess the impact of comorbidity on severity in the derivation of the PMC Severity Scale. The PMC system identifies the variety of comorbid combinations that are present, and the PMC Severity Scale quantifies the additive, hierarchical, or interactive impact of each patient's specific combination of diseases and complications. That is, in the PMC system, whether a patient with comorbid conditions is more severely ill than a single-disease patient depends on the patient's specific comorbid combination.

In addition to the clinical basis of the PMCs and related content validity of the PMC Severity Scale, a number of empirical analyses have been conducted to test the validity of the PMC Severity Scale and its reliability over time and across geographic regions. To test the construct validity of the resultant scale, the death rate and average length of stay were examined by severity level for all patients in the database used for development (Maryland 1989). Mortality rates and length of stay (as an indirect measure of morbidity) are frequently used as surrogates of severity in large population databases. Table 3 illustrates the direct relationship of both mortality rates and average lengths of stay with the severity scale using the developmental 1989 Maryland database. Results of the Kruskal Wallis $H$-test indicate a significant difference ($p < .01$) in both LOS and mortality across severity levels.

The validity of the PMC Severity Scale and its reliability over time and across geographic regions was also assessed using three databases other than the one used in scale development. First the methodology was applied to two additional years of Maryland statewide patient discharge data (1988,

Table 3: Average Mortality Rates and Lengths of Stay by Severity Level Developmental Database–Maryland 1989

| *Severity Level* | *Number of Patients* | *Mortality Rate* | *Average LOS (Discharged Alive)* | *s.d.* |
|---|---|---|---|---|
| 1 | 255,742 | 0.2 | 3.5 | 1.8 |
| 2 | 47,150 | 1.1 | 5.7 | 3.3 |
| 3 | 70,637 | 1.8 | 7.3 | 5.0 |
| 4 | 30,614 | 4.6 | 9.4 | 6.0 |
| 5 | 20,640 | 10.9 | 11.5 | 7.2 |
| 6 | 13,398 | 22.8 | 15.5 | 10.8 |
| 7 | 2,523 | 47.2 | 30.0 | 22.9 |

$n = 576{,}659$; and 1990, $n = 593{,}485$) and then to a six-month (January–June 1990) patient database of approximately 1.8 million cases from California. Patients transferred to an acute care facility or discharged against medical advice, LOS outliers, and suicide victims were excluded from the analyses.

As shown in Table 4, both mortality rates and average LOS are directly related to the PMC Severity Scale in all three validation databases. As expected, within each database, the relative ranking of severity is similar with respect to both measures, increasing with higher levels of severity. This result occurs even though, in the highest severity level, California has a much shorter length of stay as well as a higher mortality rate. Also as expected, in each severity level, Maryland's average lengths of stay and mortality rates have decreased over time (1988 to 1990). Within each population, the death rate associated with each level is approximately twice that of the previous level, indicating that the PMC Severity Scale, with further testing, may be appropriately used as an interval scale.

For each year's data, the Kruskal Wallis $H$-test and a test of pairwise comparisons were used to compare LOS and mortality distributions across severity levels. Results indicate that the differences in mortality rates and LOS among the seven levels of severity are significant ($p < .01$). By contrast, in assessing the reliability of the LOS and mortality distributions across years

Table 4: Average Mortality Rates and Lengths of Stay by Severity Level Validation Databases

| | *Average Mortality Rates* | | | *Average Lengths of Stay* | | |
|---|---|---|---|---|---|---|
| *Severity Level* | *Maryland 1988* | *Maryland 1990* | *California 1990* | *Maryland 1988 (s.d.)* | *Maryland 1990 (s.d.)* | *California 1990 (s.d.)* |
| 1 | 0.2 | 0.2 | 0.2 | 3.8 (2.2) | 3.4 (1.7) | 3.1 (1.5) |
| 2 | 1.1 | 1.0 | 0.9 | 6.0 (3.7) | 5.5 (3.3) | 4.6 (2.6) |
| 3 | 1.9 | 1.8 | 1.7 | 7.5 (5.0) | 6.8 (4.5) | 6.2 (4.2) |
| 4 | 4.8 | 4.4 | 4.2 | 9.9 (6.4) | 8.9 (5.7) | 7.8 (5.1) |
| 5 | 11.6 | 10.0 | 10.6 | 11.7 (7.2) | 11.0 (6.8) | 9.0 (5.8) |
| 6 | 24.6 | 20.9 | 22.1 | 15.9 (11.0) | 14.8 (10.2) | 12.0 (8.6) |
| 7 | 51.1 | 43.8 | 56.2 | 32.0 (25.1) | 31.1 (23.9) | 18.6 (13.0) |

(of Maryland data), the same statistical tests showed no significant difference ($p < .01$).

In order to assess the effectiveness of the PMC Severity Scale in independently predicting the outcome variables, LOS and mortality, separate regression models were estimated within each population. Linear regressions were performed to explain the variation in LOS while logistic regressions were performed to assess the relationship between each of the severity levels and the probability of survival.

Results from the linear regressions showed that, after adjusting for other independent variables (limited to age, select complication PMCs, and intensity of treatment requirements), the severity levels explained approximately 44 percent of the variation in LOS in 1989. Similarly adjusted $R^2$ values of approximately .41, .44, and .38 were found using data from Maryland 1988, Maryland 1990, and California 1990, respectively. As hypothesized, the models indicated a direct relationship between LOS and the severity levels; that is, as severity levels increased, LOS increased as well.

Overall, the PMC Severity Scale had a similar impact on LOS across the three years of Maryland data as indicated by the less than 1 percent difference in the models' parameter estimates. Also, within each year, the estimates increased in magnitude with each higher level of severity, with Level 7 having the greatest effect on LOS. The California data had similar results, differing by less than 1 percent in its estimates from that of Maryland, while showing an increasing impact of each severity level on LOS. To assess the effect of PMC severity levels on mortality, logistic regressions were performed within each population with survival (alive or dead at discharge) as the dependent variable, and the severity levels as independent dichotomous variables (along with other risk factors such as age, acuity at admission, select complication PMCs, and intensity of treatment). Two measures used to measure the "goodness of model fit" for a logistic regression are the *R*- and *C*-statistics. The *R*-statistic is similar to the multiple correlation *R* used in OLS (ordinary least squares), which adjusts for the number of parameters, while the *C*-statistic represents the area under the ROC (receiver operating characteristic) curve. The *R*-statistic ranges in value from $-1$ to $+1$; the *C*-statistic, while attaining a value of .50 through random prediction, may achieve a maximum value of 1.0. For both statistics, values closer to 1.0 are indicative of greater association to the dependent variable and, consequently, better predictability and overall fit.

Values of the *R*-statistic within each year of the Maryland data, ranging from .62 to .67, indicate a positive correlation between the PMC severity levels and inpatient mortality. The strength of this relationship is similar across years, and to that found by using the California database (.65). The *C*-statistic is a measure of the ability of the severity levels to distinguish

between survivors and nonsurvivors, for a given patient's disease(s). For the Maryland databases, *C*-values ranged from .67 to .71 while the *C*-value was .76 for the California database. Thus, within each population, the severity levels were able to discriminate between those patients expected to survive versus those expected to die for a given decision threshold.

To determine the reliability of the severity levels in predicting mortality rates within each year of data, estimated and observed rates (deaths/100 discharges) were calculated and then tested for significant differences. Using a *Z*-statistic, results indicated no significant differences ($p < .01$), either for the entire year of data or stratified by severity level within the year.

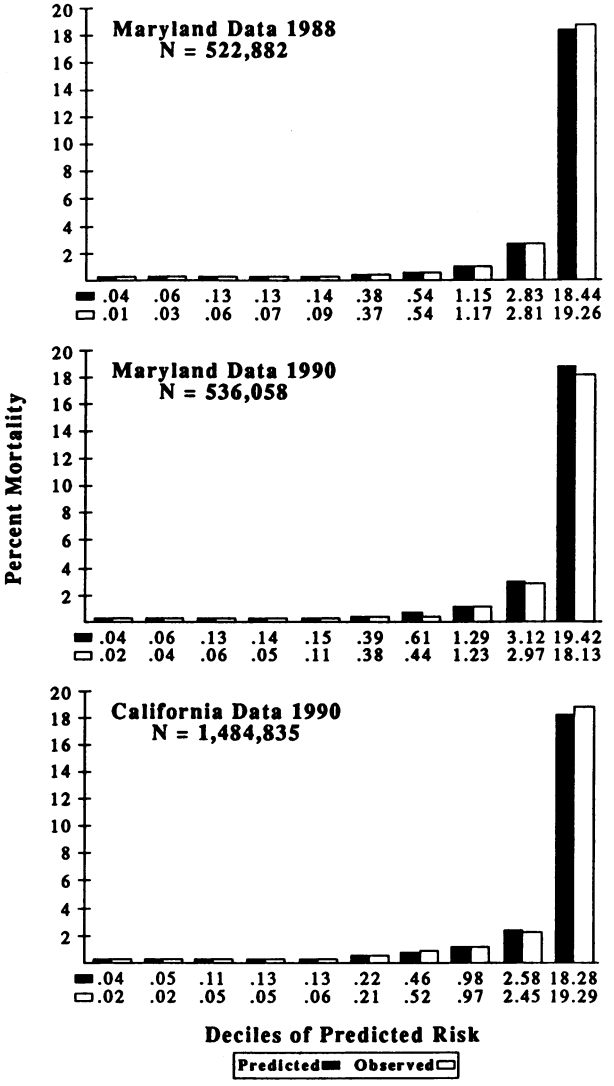
The calibration of the mortality model with observed death rates was assessed by ranking patients by their predicted probabilities of death and then dividing the data into ten equal groups. The comparison between predicted and observed mortality rates across these deciles was calculated using the Hosmer-Lemeshow chi-square statistic (Hosmer and Lemeshow 1989). Figure 2 illustrates the range of predicted and observed values for the Maryland 1988 and 1990 databases, as well as the California 1990 database. The predicted mortality rates did not vary significantly from the observed mortality rates across these deciles of risk (1988 MD: chi-square = 80.0, $p = .24$; 1990 MD: chi-square = 90.0, $p = .23$; 1990 CA: chi-square = 62.5, $p = .26$). As expected, the distribution of death rates in all three databases is highly skewed given the large number of hospitalized patients who have uncomplicated, manageable illnesses. Among these patients, death is, and is predicted to be, an infrequent event.

These results show that the independent relationships between the PMC Severity Scale with mortality and with length of stay are statistically different across levels within each year, but the relationships are statistically similar over time. Furthermore, the PMC Severity Scale was determined to be a stable predictor of mortality and LOS across two diverse geographic regions.

## SUMMARY AND DISCUSSION

The objective of this study has been to develop a direct method of assessing the relative severity of patients managed in acute care settings–one that can be used across diseases and for patients with comorbidity and/or other complicating condition(s). Patient Management Categories provided the framework used for constructing a seven-point PMC Severity Scale to quantify the severity of each patient's clinical condition over the course of a hospitalization, including the impact of coexisting conditions and complications on both the length of the hospital stay and mortality. As a result,

Figure 2: Observed and Predicted Mortality Rates for Hospitalized Patients in Three Test Databases




each patient in any large database that is consistent with Uniform Hospital Discharge Data Set (UHDDS) guidelines can be assigned one overall severity score, representing the patient's combined risk of morbidity and mortality.

The PMC Severity Scale was derived and tested on very large databases from distinct time periods and geographic regions of the United States. However, a severity index by itself–even if it is both internally and externally valid as, this one is–is not sufficient for the kind of comprehensive analyses required to assess health care outcomes and treatment effectiveness. Rather, it is critical that such a severity scale be used in conjunction with accurate case-finding methods as well as other analytic tools that measure different aspects of the care rendered.

Because of the clinical basis of PMCs, the existence of Patient Management PATHs, and the associated scalar measurements of both resource needs (PMC-RIS) and severity (PMC Severity Scale), this patient classification has wide applicability in areas such as hospital management, quality assessment, and utilization monitoring. The integrated parts of this system offer physicians, hospital managers, payers, and health services researchers a methodology for monitoring and evaluating the effectiveness of health care delivery and financing systems. PMCs and the PMC Relative Intensity Score have been used as the basis for hospital payment, but they also can be used effectively to monitor case-mix complexity and quality indicators, to manage the allocation of resources among clinical programs within hospitals, and to assess provider performance with respect to costs and utilization.

With the development of the PMC Severity Scale, an additional independent measure is available to assess patient care outcomes and the effectiveness of various treatment modalities within a set of heterogenous patients. Typically, such measures are used either as a screening mechanism to identify individual cases for more extensive clinical review or as a way of standardizing the population for severity differences in outcome studies (e.g., length of stay and death rate comparisons, risk modeling, development of treatment protocols). Using a single hospital's patient population, most patient categories will have very low frequencies. Because of these small numbers, statistical measures of severity, and sometimes costs, are used to aggregate patients, thus achieving enough patients for the performance of various types of analyses.

It should be noted that, although the PMC Severity Scale has been validated across diseases in large databases of hospitalized patients, it has not yet been applied or tested within diseases, except in selected research studies (Young, Macioce, and Young 1990; Young, Young, Smith, et al. 1991). Similarly, because the mortality assessments included in the derivation of the scale were limited to in-hospital deaths, caution should be used when applying this scale within disease areas that may have higher mortality when postdischarge data are available. Such data, however, are typically not available except for Medicare beneficiaries. Finally, underlying the PMC

Severity Scale is a diagnosis-based patient classification, with all of the limitations and benefits of ICD-9-CM coding and large databases.

Nevertheless, what is unique about the PMC Severity Scale and the underlying classification system is the specific identification of coexisting diseases and conditions that complicate patient management. Using the methods described here, all relevant comorbidity and/or specific complications are identified first. Explicit adjustments are then made for the increased resource requirements (PMC-RIS) and the increased severity (PMC Severity Scale) associated with each patient. Although severity distinctions must be made among categories in order to achieve accurate patient classification, the result of the first step is still a nominal classification. The derivation of the PMC Severity Scale, along with the prior derivation of the PMC Relative Intensity Score, permits the independent evaluation of patient severity versus intensity of resource requirements (costliness) for large populations of hospitalized patients. Thus, the PMC Severity Scale will make it easier to analyze patient outcomes in large databases and to examine, for particular diseases, the extent to which severity and intensity are related.

## NOTES

1. The PMC computerized classification is licensed to users by The Pittsburgh Research Institute, a 501(c)(3) not-for-profit health services research organization. License fees are used to support the clinical and technical research required to keep the PMC system consistent with changes in health care practice and technology. Neither the authors, nor any other individuals, have proprietary interests in PMCs. That is, no personal financial gain is associated with the dissemination of PMCs.
2. The variables used by the PMC classification software include up to ten ICD-9-CM diagnosis codes and ten ICD-9-CM procedures, along with age and gender of the patient. These variables, as well as other core data elements such as admission and discharge dates, disposition of the patient (e.g., discharged home alive, died, transferred), are included in the Uniform Hospital Discharge Data Set (UHDDS), which is a minimum basic data set collected by virtually all acute care hospitals in the United States. Most electronically submitted hospital insurance claims (Medicare, Medicaid, Blue Cross) are in the form of UB-82 (Uniform Bill, 1982), which is consistent with UHDDS guidelines. In some instances, hospital data and third party payer data are aggregated and maintained by statewide agencies (which was the source of the databases used in this study).
3. The PMC Classification System has been applied to the following statewide databases: Florida, Iowa, Maine, Maryland, New York, Pennsylvania, and Washington. PMCs have also been the basis of projects in Spain, Australia, the United Kingdom, and Germany.

4. The relationship between severity of the patient's illness at a point in time and management difficulty has been confirmed recently, but only among a small number of resident physicians (Kelleher 1993).
5. The combinatorial formula $[{}_nC_r = n!/r!(n\text{-}r)!]$ calculates the number of different combinations of size $r = 1, 2, 3$, and 4, that can be formed from $n = 4$ possible severity levels, resulting in the following 15 severity level combinations:

| *Notation* | *Combinations* |
|---|---|
| ${}_4C_1$ | {1}, {2}, {3}, {4} |
| ${}_4C_2$ | {1&2}, {1&3}, {1&4}, {2&3}, {2&4}, {3&4} |
| ${}_4C_3$ | {1&2&3}, {1&2&4}, {1&3&4}, {2&3&4} |
| ${}_4C_4$ | {1&2&3&4} |

6. Because comorbid cases with four or more PMCs represented by the same severity level accounted for less than 10 percent of the data, these cases were combined into a "3+" severity level.

## REFERENCES

Charbonneau, C., C. Ostrowski, E. T. Poehner, P. Lindsay, T. L. Panniers, P. Houghton, and J. Albright. "Validity and Reliability Issues in Alternative Patient Classification Systems." *Medical Care* 26, no. 8 (1988): 800–13.

Charlson, M. E., P. Pompei, K. L. Ales, and C. R. MacKensie. "A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation." *Journal of Chronic Disease* 40, no. 5 (1987): 373–83.

Dubois, R. W., J. H. Moxley, D. Draper, and R. H. Brook. "Hospital Inpatient Mortality: Is It a Predictor of Quality?" *New England Journal of Medicine* 317, no. 26 (1987): 1674–80.

Gonnella, J. S., M. C. Hornbrook, and D. Z. Louis. "Staging of Disease: A Case-Mix Measurement." *Journal of the American Medical Association* 251, no. 5 (1984): 637–44.

Greenfield, S., H. U. Aronow, R. M. Elashoff, and D. Watanabe. "Flaws in Mortality Data: The Hazards of Ignoring Comorbid Disease." *Journal of the American Medical Association* 260, no. 15 (1988): 2253-55.

Hosmer, D. W., and S. Lemeshow. *Applied Logistic Regression.* New York: John Wiley & Sons, 1989: pp. 140–45.

Hospital Research and Educational Trust. *The Hospital Administrator's Guide to Severity Measurement Systems.* Chicago: American Hospital Association, 1989.

Kelleher, C. "Relationship of Physician Ratings of Severity of Illness and Difficulty of Clinical Management to Length of Stay." *Health Services Research* 27, no. 6 (February 1993): 841–55.

Stein, R. E., S. L. Gortmaker, E. C. Perrin, I. B. Pless, D. K. Walker, and M. Weitzman. "Severity of Illness: Concepts and Measurements." *The Lancet* 2, no. 8574 (1987): 1506–09.

Thomas, J. W., J. J. Holloway, and K. E. Guire. "Validating Risk-Adjusted Mortality as an Indicator for Quality of Care." *Inquiry* 30 (Spring 1993): 6–22.

Thomas, J. W., and D. R. Longo. "Application of Severity Measurement Systems for Hospital Quality Management." *Hospital & Health Services Administration* 35, no. 2 (Summer 1990): 221–43.

Young, J. C., D. P. Macioce, and W. W. Young. "Identifying Injuries and Trauma Severity in Large Databases." *The Journal of Trauma* 30, no. 10 (October 1990): 1220–30.

Young, W. W., S. Kohler, and D. P. Macioce. *Patient Management Categories, Release 5.0, Volumes I and II.* Pittsburgh, PA: The Pittsburgh Research Institute, July 1992.

Young, W. W. "Incorporating Severity of Illness and Comorbidity in Case-Mix Measurement." *Health Care Financing Review Annual Supplement* 6 (November 1984): 23–31.

Young, W. W., D. Z. Joyce, J. L. Schuchert, M. A. Hutton, and D. P. Macioce. *Measuring the Costs of Care Using Patient Management Categories, Final Report. Volume I: Patient Management Categories; Volume II: PMC Relative Cost Weights; and Volume III: Patient Management PATHs.* HCFA Grant Number 18-P-97063/3, May 1985. Available through the National Technical Information Service: Volume I, NTIS No. PB86101979; Volume II, NTIS No. PB86101987.

Young, W. W., J. C. Young, J. S. Smith, and M. Rhodes. "Defining the Major Trauma Patient and Trauma Severity." *The Journal of Trauma* 31, no. 8 (August 1991): 1125–41.