Reprinted from ANNALS OF INTERNAL MEDICINE Vol. 118; No. 10, 15 May 1993
Printed in U.S.A.

Systems for Comparing Actual and Predicted Mortality Rates: Characteristics To Promote Cooperation in Improving Hospital Care

During the past decade, mortality rates have been increasingly proposed as a measure of quality of hospital care. In 1986, the federal government released hospital mortality rates for Medicare patients in order to allow hospitals and consumers to evaluate care (1, 2). These now-yearly releases include each hospital's mortality rates for 17 broad diagnostic groups, along with the *expected* mortality rates based on the group's severity of illness as assessed from Medicare reimbursement claims. For example, if a hospital's cardiac patient group has a severity-adjusted actual mortality rate of 15%, then a predicted rate of 5% could suggest suboptimal care, whereas a predicted rate of 30% would reflect outstanding care. Such comparisons of actual and severity-adjusted predicted mortality rates should facilitate hospitals' assessment of their own performance as well as permit fair comparisons among hospitals.

Unfortunately, because the Medicare predicted mortality rates have very limited adjustments for severity of illness within given diagnostic groups, it is difficult to say that differences between a given hospital's actual and predicted mortality rates are in fact related to quality of care (3-5). Indeed, there is consensus in the health care industry and among researchers that if such data are to be used, they must be accurately risk adjusted. Over the years, several other mortality predictors have been developed that include far more detailed severity adjustments (6); the performance of one of these, APACHE III, the third in the series of Acute Physiologic, Age, Chronic Health Evaluation (APACHE) scores, is the subject of an article (7) in this issue. Such tools have brought the possibility of using risk-adjusted mortality rates to assess quality of hospital care closer to reality.

Indeed, despite many concerns (1, 2, 8-10), there are attractive reasons for using mortality rates as one measure of quality of medical care. Mortality is an objective end point and generally is more uniformly available than many other potential quality-of-care indicators. Moreover, for certain conditions, it measures a very relevant outcome. For example, severity-adjusted mortality rates are useful in assessing quality of care for patients with acute myocardial infarction but not in assessing terminal care for patients with metastatic cancer. The new work by the APACHE group (7) shows the power of using a risk-adjusted mortality predictor to assess medical care outcomes as a measure of quality of care. As experience with mortality predictor systems increases (6, 8, 9, 11), it is becoming clearer what characteristics such systems should have.

By themselves, mortality rates, whether used alone or in conjunction with severity-adjusted predicted rates, will not improve care. Only an appropriate response to mortality rates by clinicians, hospital administrators, consumers, and health care payers will improve care. Thus, if severity-adjusted mortality predictors are to have their desired effect, they must be attractive to, and promote cooperation among, these groups. The specific characteristics that mortality predictor systems must have to achieve this goal are listed below.

1. To be useful in patient management, and thus to be trusted by clinicians, a risk-adjusted mortality predictor should be sufficiently accurate and easy to use in real-time clinical settings. Also, it should provide the same mortality prediction when used prospectively in clinical care as when used retrospectively; that is, it should be a "time-insensitive predictive instrument" (TIPI) (11, 12).

2. To accurately capture a patient's true *presenting* risk for death (and to allow use in the real-time clinical setting), mortality should be predicted based on the first minutes of hospital presentation, not based on data from the first full 24 hours or more after presentation, as is the case with APACHE and most other mortality predictors. Severity of illness within the first 24 hours may be more a measure of suboptimal care than severity of initial presentation. This is especially so for acute cardiac, pulmonary, and surgical problems, for which interventions within the first minutes and hours are extremely important and may account for a large proportion of preventable deaths within the first 24 hours.

3. A mortality predictor should not be affected by whether a patient is hospitalized. Some patients presenting to hospitals do not generate hospital records because they are triaged directly to home or to another institution, or because they do not survive long enough to be admitted. Depending on the hospital type, its aims, its community role, and the quality of its emergency care, the likelihood of admission after the initial evaluation may vary widely. Thus, a mortality predictor that only captures a presenting patient if an inpatient record is generated may be significantly biased and misleading.

4. Mortality predictor systems should require only data that are collected in the usual care of the patients. If more weight is given to high-technology or special tests (for example, the MedisGroup system (Mediqual,

Westborough, Massachusetts) that gave more severity credit to ascites detected by sonogram than to ascites determined by physical examination [13]), then hospitals will be rewarded for using more tests, regardless of appropriateness or quality of care.

5. Even if only basic clinical data are used, the dependence of a mortality system on medical record documentation can introduce biases. The quantity and quality of documentation vary among different types of hospitals and are related to factors that also affect real care and outcomes: differences in patient populations due to referrals, physician practice types, the availability of special expertise and facilities, and the use of physician trainees. Differences in the use of facilities may also create biases; for example, the more detailed data generated in an intensive care unit are likely to yield a higher severity score than the data collected for the same patient on a regular ward. Thus, if a mortality predictor system depends on the medical record, it only should use objective and universally collected information.

6. Risk-adjusted mortality predictors should not depend on or be sensitive to hospital reimbursement claims–based diagnoses. Assignment of diagnoses can be biased by individual and local practice patterns, by the need to use certain diagnoses to gain access to special resources in the acute care setting (for example, the use of the diagnosis "rule-out myocardial infarction" to gain access to a coronary care unit bed), and especially by the economic incentive to assign more highly reimbursed diagnoses. Indeed, Iezzoni and colleagues (14) showed that there was as high as 40% overcoding of the ICD-9 acute myocardial infarction diagnosis in hospitals' claims, presumably related to reimbursement implications. A large admixture of the less severely ill patients *without* acute infarction yields a group with a lower mortality rate than if only the patients with true acute infarctions were included. Thus, patient inclusion criteria for risk-adjusted mortality predictors should define comparable patients in a clinically meaningful group not susceptible to biases or differences in practice settings.

7. A mortality system's predictions must be accurate at all levels of its severity scale; that is, it must have excellent "calibration." Indeed, the regression techniques on which most mortality predictors are based have a propensity for poor calibration; that is, they tend to underpredict the likelihood of death for more severely ill patients and overpredict the likelihood of death for patients with less severe illness. This creates a pernicious bias: 1) Hospitals providing care to more severely ill patients will tend to have actual mortality rates above predicted and thus will appear to be giving poor care; and 2) hospitals providing care to less severely ill patients will tend to have actual mortality rates lower than predicted and thus will appear to be giving better than average care. Thus, a poorly calibrated system will reward hospitals that successfully avoid providing care to more severely ill patients, which will merely reinforce the already existing perverse incentives of fixed Diagnosis Related Group (DRG)–based payment to avoid the care of the sickest patients. Unfortunately, no data are available on the calibration of most mortality predictor systems (6). Calibration data have been reported for the APACHE systems (7, 15, 16), the acute myocardial infarction mortality time-insensitive predictive instrument (11), the Medicare Mortality Predictor System (17), and the Mortality Prediction Model (18). Such data should be reported for all systems (6, 19, 20).

8. Given that hospitals have mushrooming numbers of mandated medical record reviews and reports, a mortality predictor system must not add substantially to this burden. Severity systems should require only commonly available equipment and should be compatible with other computerized clinical data systems in general use. For example, APACHE II and III as well as other systems are available on computer-based clinical information systems, and the acute infarction mortality time-insensitive predictive instrument can be implemented in a computerized electrocardiograph that will automatically generate and store its predictions (11). By such computer system integration, risk-adjusted mortality comparisons can be done entirely without medical records, representing a substantial advance in hospital operations and medical care.

9. To be trusted and accepted by all those whose performance will be directly or indirectly judged by a mortality-based severity system, the components of the system must be open for inspection and testing. Risk-adjusted mortality predictors are becoming increasingly important to the health care system, but distrust could remain a major hindrance to their use (3, 4, 20). The APACHE I and II (15, 16) and work by other groups (7, 11, 17-19) have provided such details, and one hopes that the APACHE III owners (7) and others (21, 22) will soon follow suit.

As a prototype mortality predictor system, the APACHE III meets about half of the criteria listed above. The predictive performance and calibration of the APACHE II (15, 16) and the APACHE III (7) are excellent. Commercial products made by the authors' company and other companies can provide APACHE scores via computerized information systems, although only for hospitalized patients. The APACHE systems have characteristics that are likely to be familiar and attractive to clinicians. However, it is important to note that APACHE scores are based on the first 24 hours of care rather than on only the first few minutes and that they have not yet been shown to be effective or safe in a prospective trial as a clinical decision aid (23). Finally, the APACHE investigators have declined to make the actual formulae of the APACHE III public, and currently these formulae can only be evaluated with restrictions specified by APACHE Medical Systems, Inc. Most other currently available severity systems also generally fail to meet some or most of these criteria, although a system meeting all these criteria has been demonstrated (11, 12).

The task remains to develop and refine severity-adjusted mortality systems that meet all these criteria. Not only will such systems provide the risk-adjusted predictions now in such great demand, but they also will promote cooperation among clinicians, hospitals, health care systems, and payers and thereby greatly enhance chances for the actual improvement of medical care.

Harry P. Selker, MD, MSPH
New England Medical Center
Tufts University School of Medicine
Boston, MA 02111

Requests for Reprints: Harry P. Selker, MD, MSPH, Center for Cardiovascular Health Services Research, New England Medical Center, 1031, 750 Washington Street, Boston, MA 02111.

Annals of Internal Medicine. 1993;**118:**820-822.

References

1. **U.S. Health Care Financing Administration.** Medicare hospital mortality information. Washington, DC; Government Printing Office; 1987.
2. **U.S. Health Care Financing Administration.** Medicare hospital mortality information. Washington, DC; Government Printing Office; 1988.
3. **Green J, Wintfeld N, Sharkey P, Passman LJ.** The importance of severity of illness in assessing hospital mortality. JAMA. 1990;263: 241-6.
4. **Berwick DM, Wald DL.** Hospital leaders' opinions of the HCFA mortality data. JAMA. 1990;263:247-9.
5. **Park RE, Brook RH, Kosecoff J, Keesey J, Rubenstein L, Keeler E, et al.** Explaining variations in hospital death rates. Randomness, severity of illness, quality of care. JAMA. 1990;264:484-90.
6. **Hadorn DC, Keiler EB, Brook RH, Rogers WB.** Assessing the performance of mortality predictive models. Santa Monica, California: Rand; 1993. RAND Publication # MR-181-HCFA.
7. **Knaus WA, Wagner DP, Zimmerman JE, Draper EA.** Variations in hospital mortality and length of stay from intensive care. Ann Intern Med. 1993;118:753-761.
8. **Knaus WA, Nash DB.** Predicting and evaluating patient outcomes [Editorial]. Ann Intern Med. 1988;109:521-2.
9. **Kahn KL, Brook RH, Draper D, Keeler EB, Rubenstein LV, Rogers WH, et al.** Interpreting hospital mortality data. How can we proceed? JAMA. 1988;24:3625-8.
10. **Greenfield S, Aronow HU, Elashoff RM, Watanabe D.** Flaws in mortality data. The hazards of ignoring comorbid disease. JAMA. 1988;260:2253-5.
11. **Selker HP, Griffith JL, D'Agostino RB.** A time-insensitive predictive instrument for acute myocardial infarction mortality: a multicenter study. Med Care. 1991;29:1196-211.
12. **Selker HP, Griffith JL, D'Agostino RB.** A tool for judging coronary care unit admission appropriateness valid for both real-time and retrospective use. A time-insensitive predictive instrument (TIPI) for acute cardiac ischemia: a multicenter study. Med Care. 1991; 29:610-27.
13. **Iezzoni LI, Moskowitz MA.** A clinical assessment of MedisGroups. JAMA. 1988;260:3159-63.
14. **Iezzoni LI, Burnside S, Sickles L, Moskowitz MA, Sawitz E, Levine PA.** Coding of acute myocardial infarction. Clinical and policy implications. Ann Intern Med. 1988;109:745-51.
15. **Knaus WA, Draper EA, Wagner DP, et al.** APACHE II: A severity of disease classification system. Crit Care Med. 1985;13:818-29.
16. **Knaus WA, Draper EA, Wagner DP, Zimmerman JE.** An evaluation of outcome from intensive care in major medical centers. Ann Intern Med. 1986;104:410-8.
17. **Daley J, Jencks SF, Draper D, Lenhart G, Thomas N, Walker J.** Predicting hospital-associated mortality for medicare patients with stroke, pneumonia, acute myocardial infarction, and congestive heart failure. JAMA. 1988;260:3617-24.
18. **Teres D, Lemeshow S, Avrunin SJ, Pastides H.** Vālidation of the mortality prediction model for ICU patients. Crit Care Med. 1987; 15:208-13.
19. **Jencks SF, Daley J, Draper D, Thomas N, Lenhart G, Walker J.** Interpreting hospital mortality data: the role of clinical risk adjustment. JAMA. 1988;260:3611-6.
20. **Schroeder SA.** Outcome assessment 70 years later: are we ready? [Editorial]. N Engl J Med. 1987;316:160-2.
21. **Brewster AC, Karlin BG, Hyde LA, Jacobs CM, Bradbury RC, Chae YM.** MEDISGRPS: A clinically based approach to classifying hospital patients at admission. Inquiry. 1985;12:377-87.
22. **Horn SD, Horn RA.** The Computerized Severity Index. A new tool for case-mix management. J Med Syst. 1986;10:73-78.
23. **Selker HP.** Coronary care unit triage decision aids: how do we know when they work? [Editorial]. Am J Med. 1989;87:491-93.