The SUPPORT Prognostic Model. Objective Estimates of Survival for Seriously Ill Hospitalized Adults. Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatme...

Article *in* Annals of Internal Medicine · March 1995

Source: PubMed

CITATIONS
244

READS
667

15 authors, including:

William A. Knaus
University of Virginia
337 PUBLICATIONS **73,086** CITATIONS
SEE PROFILE

Frank E Harrell
Vanderbilt University
450 PUBLICATIONS **90,105** CITATIONS
SEE PROFILE

Joanne Lynn
Altarum Institute
423 PUBLICATIONS **34,127** CITATIONS
SEE PROFILE

Russell S Phillips
Harvard Medical School
405 PUBLICATIONS **40,403** CITATIONS
SEE PROFILE

The SUPPORT Prognostic Model

Objective Estimates of Survival for Seriously Ill Hospitalized Adults

William A. Knaus, MD; Frank E. Harrell Jr., PhD; Joanne Lynn, MD, MA; Lee Goldman, MD, MPH; Russell S. Phillips, MD; Alfred F. Connors Jr., MD; Neal V. Dawson, MD; William J. Fulkerson Jr., MD; Robert M. Califf, MD; Norman Desbiens, MD; Peter Layde, MD, MSc; Robert K. Oye, MD; Paul E. Bellamy, MD; Rosemarie B. Hakim, PhD; and Douglas P. Wagner, PhD

■ *Objective:* To develop and validate a prognostic model that estimates survival over a 180-day period for seriously ill hospitalized adults (phase I of SUPPORT [Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments]) and to compare this model's predictions with those of an existing prognostic system and with physicians' independent estimates (SUPPORT phase II).

■ *Design:* Prospective cohort study.

■ *Setting:* 5 tertiary care academic centers in the United States.

■ *Participants:* 4301 hospitalized adults were selected for phase I according to diagnosis and severity of illness; 4028 patients were evaluated from phase II.

■ *Measurements:* A survival model was developed using the following predictor variables: diagnosis, age, number of days in the hospital before study entry, presence of cancer, neurologic function, and 11 physiologic measures recorded on day 3 after study entry. Physicians were interviewed on day 3. Patients were followed for survival for 180 days after study entry.

■ *Results:* The area under the receiver-operating characteristics (ROC) curve for prediction of surviving 180 days was 0.79 in phase I, 0.78 in the phase II independent validation, and 0.78 when the acute physiology score from the APACHE (Acute Physiology, Age, Chronic Health Evaluation) III prognostic scoring system was substituted for the SUPPORT physiology score. For phase II patients, the SUPPORT model had equal discrimination and slightly improved calibration compared with physicians' estimates. Combining the SUPPORT model with physicians' estimates improved both predictive accuracy (ROC curve area = 0.82) and the ability to identify patients with high probabilities of survival or death.

■ *Conclusions:* A limited amount of readily available clinical information can provide a foundation for long-term survival estimates that are as accurate as physicians' estimates. The best survival estimates combine an objective prognosis with a physician's clinical estimate.

Ann Intern Med. 1995;122:191-203.

From the George Washington University Medical Center, Washington, D.C.; Duke University Medical Center, Durham, North Carolina; Dartmouth Medical School, Hanover, New Hampshire; Beth Israel Hospital, Boston, Massachusetts; Metrohealth Medical Center, Cleveland, Ohio; Marshfield Medical Research Foundation and the Marshfield Clinic, Marshfield, Wisconsin; and University of California, Los Angeles, School of Medicine, Los Angeles, California. For current author addresses, see end of text.

The Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments (SUPPORT) was a multicenter study designed to examine outcomes and clinical decision making for seriously ill hospitalized patients (1). A major hypothesis of SUPPORT was that accurate prediction of risk for death might assist physicians in clinical decision making by decreasing uncertainty and by promoting communication among physicians, patients, and patients' families (2, 3). SUPPORT was designed to be completed in two phases. During phase I, we observed and described the natural history of decision making and developed models to predict outcomes. Phase II was a randomized clinical intervention trial in which we evaluated providing objective prognostic information and enhanced communication about prognoses and preferences.

For phase I, we enrolled all patients at five participating sites who had at least one of nine illnesses and who were expected to have an overall 6-month mortality of 50% (Appendix 1). Individual survival probability estimates were developed for these patients using a few readily available variables that described the patient's major disease class, severity of physiologic abnormality, age, and comorbid conditions. We attempted to improve on past prognostic efforts in three ways. First, the SUPPORT population included patients who were not severely physiologically imbalanced, such as patients treated outside of intensive care units; most previous prognostic systems have been confined to intensive care units or emergency rooms (4–7). Second, the SUPPORT model was designed to predict survival to 180 days after study entry rather than to hospital discharge. Third, the independent variables used to predict risk for death were allowed to assume nonlinear relations that more accurately reflected their biological relations with patient survival. We summarize these efforts by describing the development, performance, and validation of the SUPPORT prognostic model, and we compare the model both with previous efforts and with the simultaneous subjective prognostic estimates of the physicians who cared for the study patients.

Methods

Patient Selection

A literature review (8) identified 13 diagnostic groups that had sufficient prognostic information in the medical record to allow identification of a cohort of patients with an aggregate expected 180-day mortality rate of 50%. Pilot testing eliminated 4 of these groups because of inadequate sample size, unreliable estimation of staging from the chart, or relatively low subsequent mortality

rate (8). Patients in each of the remaining 9 groups (Appendix 1) had to be more than 18 years of age and were excluded if they died within 48 hours of hospitalization or if they were scheduled for discharge within 72 hours of admission. Patients were also excluded if they had the acquired immunodeficiency syndrome, were admitted with head trauma, were pregnant, had trauma other than acute respiratory failure or multiple organ system failure, had acute burns, were admitted to the psychiatric unit, or did not speak English. Appendix 1 describes the case selection process, but additional discussion can be found in the published study design (1).

Data Collection

Phase I data were collected from June 1989 to June 1991 at Beth Israel Hospital, Boston, Massachusetts; MetroHealth Medical Center, Cleveland, Ohio; Duke University Medical Center, Durham, North Carolina; Marshfield Clinic/St. Joseph's Hospital, Marshfield, Wisconsin; and the University of California, Los Angeles, Medical Center, Los Angeles, California. Phase II data were collected from January 1992 to January 1994 at the same institutions. Data collection procedures were almost identical for both phases.

We pre-specified all variables used in the prognostic model by first developing a list of general variables that were expected to be available for all patients. These general variables included the nine diagnostic groups, physiologic variables such as vital signs (temperature, mean blood pressure, heart rate, and respiratory rate) and common laboratory measures (arterial blood gases, serum sodium, serum potassium, serum creatinine, hematocrit, leukocyte count, serum albumin, and serum bilirubin), and a clinical assessment of neurologic status done using the Glasgow coma scale. This list was based on the APACHE III prognostic classification system (4). The exact level of the most abnormal value during each specified 24-hour period and a comprehensive listing of comorbid conditions were also collected (4).

Prognostic variables specific to each of the nine diagnostic groups were also identified (1). For example, for patients with congestive heart failure, the previous ejection fraction, cardiac rhythm disturbances, history of myocardial infarction, current congestive heart failure severity (edema, rales, elevated jugular venous pressure), and new cardiac events were collected from the medical record and analyzed for their prognostic value.

The above data were collected for the first 24 hours after study entry, which was the first hospital day for all patients not in the intensive care unit. For patients in the intensive care unit, it was the first 24-hour period during hospitalization after development of acute respiratory failure, coma, or multiple organ system failure. The same variables (except for comorbid conditions) were also collected on days 3, 7, 14, and 25 after study entry. We focus on the use of day 3 data in predicting risk for death, given that all patients in our study had to survive for at least 48 hours after initial qualification for SUPPORT. Details of the models for other days are available from the authors.

Patients admitted to the hospital and those in the intensive care unit were screened daily, and all patients who met the diagnostic and severity criteria described in Appendix 1 were enrolled in the study. Data accessibility was determined by routine diagnostic and testing procedures and charting practices at the five medical centers. Study patients were followed to assess survival and functional status for 180 days after study entry. Direct 180-day follow-up was completed for 96% of phase I patients; 4% were not contacted directly, and for these we searched the 1989–1992 deaths registered in the National Death Index. Because follow-up using the National Death Index was not possible for deaths occurring after 1992, direct 180-day follow-up was used for most phase II patients.

Patient identification and data collection procedures included ongoing reliability testing. General physiologic measures were collected by a second nurse abstractor from a random 10% sample of patients within 3 to 10 days of the initial data collection. The number of hospital admissions and the number of in-hospital deaths during the study period were also obtained.

Each patient's physician—defined as the most senior physician in the health care team who could be interviewed in time and who acknowledged responsibility for decision making—was asked to give a numeric estimate of the patient's likelihood of surviving 2 and 6 months after study entry. Thirteen percent of these interviews were conducted with fellows or housestaff; all were conducted either in person or using a self-administered questionnaire and were completed between day 2 and day 6 after study entry (median day, 2.9).

Statistical Modeling

The purpose of the statistical analysis was to determine the influence of each of the pre-specified prognostic factors on estimates of an individual patient's survival time. We used a Cox proportional hazards regression model (9) after testing, with various techniques (10, 11), to be sure that this method was appropriate for the data. Because laboratory and vital signs are continuous measures but also have clinically important threshold values, we used a statistical technique called restricted cubic splines (12–14) that better approximates the true biological relation between these variables and risk for death. Restricted cubic splines allow continuous data to fit within the Cox model without assuming a linear relation (12–14). These fitted cubic splines also permit the placement of breaks in the continuous measurement of a variable if the relation between the prognostic variable and risk for death changes abruptly. For example, a mean blood pressure greater than 60 mm Hg did not contribute additional prognostic value, so the relation with mortality sloped to this point and was flat thereafter. Before attempting to simplify the representation of each variable, all terms were tested so that prognostically unimportant variables could be discarded. Variables were defined as important using Akaike's information criterion (15), which adds variables to a model until the total chi-square for all remaining variables is less than twice their aggregate degrees of freedom.

Interactions between variables and diseases were explored so that individual physiologic variables could have disease-specific relations with risk for death. Potential interactions between disease and specific prognostic variables, such as leukocyte count in patients with acute respiratory failure and albumin in patients with cancer, were pre-specified from literature reviews and past experience. The effect of age on survival in specific diagnostic classes was also examined.

To minimize bias associated with the unavailability of data in patient subgroups, a series of analyses was done to determine the most appropriate approach for imputing missing values. These analyses indicated that the most appropriate value to impute when a physiologic value was missing on day 1 of data collection was a value within the normal range. If a physiologic value was missing at day 3 of data collection, the day 1 value was used. An exception to this was serum albumin, for which we used the day 1 value throughout, substituting the day 3 value only when the day 1 value was missing.

The phase I data were used both to develop the prognostic model and to initially estimate the model's performance, which was later formally evaluated on those phase II patients for whom 180-day outcomes and physicians' estimates were available.

The area under a ROC curve was used as a measure of overall predictive discrimination, which was defined in this study as the ability to separate those patients likely to live 180 days after study entry from those likely to die before 180 days had elapsed. A ROC curve area of 0.5 indicates no discrimination and an area of 1.0 indicates perfect prediction (16). Calibration, which is the ability to predict probabilities across all ranges of risk, is illustrated with a calibration chart that plots predicted survival against observed survival. In a calibration chart, a line at a 45-degree angle represents perfect calibration.

The model's performance on study day 3 was compared with the observed survival status of patients on study day 180. Although prognostic estimates were also available for study days 1, 7, 14, and 25, study day 3 was chosen because it was the first day after study qualification when physicians' subjective estimates of survival were available. The 180-day mortality rates were chosen as the primary analytic outcome for this report.

Comparisons with the APACHE III prognostic system were made by substituting the physiologic weighting from the acute physiologic score of the APACHE III prognostic model (4) for the weighted sum of physiologic variables in the SUPPORT physiologic score and refitting the Cox model. To determine whether the survival estimates of physicians added prognostic

Table 1. Sample Size, Major Disease Groups, Demographic Characteristics, Disease Severity, and Outcome of Patients Enrolled in Phase I of SUPPORT*

Disease Group	Patients	Median Age	Men, %	Median SLOS†	Median APS‡	Deaths during Initial Hospitalization	Deaths by 180 Days after Study Entry (Kaplan-Meier)
	n	*y*	%	*d*		%	%
Acute respiratory failure	738	66	55	16	52	31	42
Chronic obstructive pulmonary disease	458	71	52	10	44	13	35
Congestive heart failure	726	68	61	8	34	6	29
Chronic liver disease	296	52	61	10	40	24	47
Coma	247	68	52	9	83	65	78
Colon cancer	269	64	55	8	16	8	45
Lung cancer	459	62	64	6	20	13	62
Multiple organ system failure with cancer	333	59	55	12	69	54	76
Multiple organ system failure with sepsis	775	62	56	19	69	42	52
Total	4301	65	57	11	47	27	48

* SUPPORT = The Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments.
† SLOS = study length of stay in initial hospitalization.
‡ Acute physiology score of APACHE III during the first 24 hours after study entry.

information to the model, we used the phase II patient sample, for which physicians' 180-day subjective estimates were available. Because Cox regression assumes that variables are linear in a log-(log) scale, the physicians' subjective estimates of patient survival were transformed with linear splines before evaluation for inclusion in the final Cox model. To further describe discrimination, we compared all cases in which a patient with a less than 15% or a greater than 85% probability for survival to 180 days was identified by any of the following four models: the SUPPORT model, the SUPPORT model with the APACHE III acute physiology score, the physician's prediction, or the SUPPORT model plus the physician's prediction. Formal tests of added prognostic information were made by fitting a Cox proportional hazards model containing two variables representing two predictions. Each variable in this model was tested with a likelihood ratio chi-square test to see whether it added independent prognostic information to the information provided by the other predictor (17).

All statistical analyses were done using the S-Plus Statistical Language, Version 3.2, and the S-Plus Reference Manual (Statistical Sciences, Inc., Seattle, Washington).

Results

Patients and Prognostic Variables

The specific entry criteria for phase I were met by 4301 patients (Appendix 1); 2072 of these patients (48%) died within 6 months of study entry (Table 1). In the five study hospitals, these patients constituted approximately 3% of admissions during the study period but accounted for approximately 19% of all in-hospital adult deaths. The patients had a median age of 63 years and 57% of them were men. Their in-hospital mortality rate ranged from 6% for those with congestive heart failure to 65% for those with coma. The 180-day mortality rates ranged from 29% for patients with congestive heart failure to 78% for those with coma (Table 1).

At the time of this report, direct 180-day follow-up was available for 4542 of the 4804 phase II patients and a physician's estimate of 180-day survival was available for 4028 of the 4804 (84%). These 4028 patients served as the basis for the validation. Characteristics of phase I and phase II patients did not differ significantly.

In the reliability study of medical record abstracting, exact agreement was 87% for specific physiologic variables and 82% for specific coexisting morbidities (18). Inter-abstractor agreement was excellent for the two most common comorbid conditions, chronic obstructive pulmonary disease and metastatic cancer ($\kappa = 0.90$).

Although the general prognostic factors were usually available in medical records, disease-specific tests were available much less often. For example, a recent ejection fraction was found in the charts of only 26% of the patients who met the study entry criteria for congestive heart failure, and pulmonary function tests were recorded by study day 3 in only 31% of patients with chronic obstructive pulmonary disease.

Derivation of the SUPPORT Prognostic Model Using Phase I Data

The SUPPORT prognostic model, like the current APACHE III model, is based on disease category, severity of acute disease as measured by physiologic abnormalities (SUPPORT physiology score), evaluation of the patient's long-term health status, and the number of days the patient was hospitalized before study entry (4). The relative importance of each major component of the model, as measured by its relative contribution to the overall chi-square, is shown in Figure 1. The higher the chi-square, the greater the prognostic significance.

Disease

The predicted risk for death varied among the nine disease groups during the follow-up period. For example, patients with coma had a high early mortality rate, but few died more than 30 days after study entry. Patients with colon or lung cancer had only half the 30-day mortality rate of patients with respiratory or multiple organ system failure, but these groups had similar 180-day mortality rates (Appendix Figure 1). A stratified Cox model was used to account for these variations in risk that violated the basic proportional hazards assumption. In the model, on the basis of similarly shaped survival curves, the nine disease groups were consolidated into four larger disease classes: acute respiratory failure and multiple organ system failure; chronic obstructive pulmonary disease,

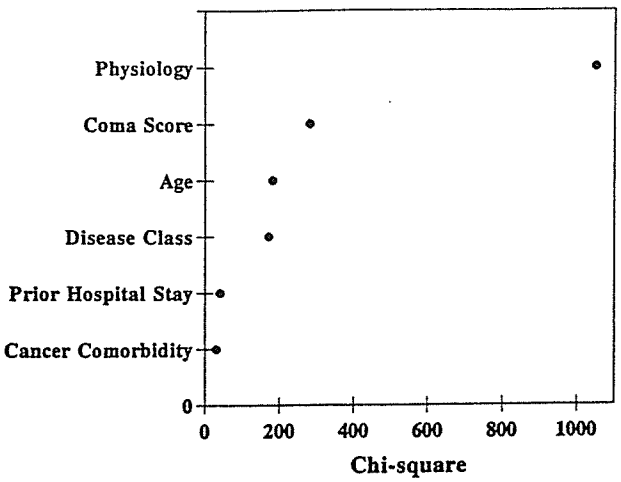


Figure 1. Relative contributions of major prognostic elements in the SUPPORT model as measured by the amount of chi-square accounted for by each element. Physiology = all physiologic measures except Glasgow coma scale; cancer comorbidity = presence of cancer in addition to disease category; previous hospital stay = days in the hospital before study entry.

congestive heart failure, and cirrhosis; coma; and colon and lung cancer. Indicator variables were used to distinguish each disease group within each of the larger disease classes.

Physiology Score: Severity of Disease

The single most important prognostic factor in the SUPPORT model was the physiology score, which was obtained from measurements done on study day 3 (Figure 1). Some factors, such as serum bilirubin levels, had a simple linear relation with survival; other variables had a nonlinear relation with survival (Appendix Figure 2). For example, mean blood pressure was associated with increased risk for death up to 60 mm Hg; a similar threshold effect was seen with PaO_2/FIO_2 (Appendix Figure 2). In another example, survival was higher, indicated by a lower risk score, in patients with serum creatinine values between 0.9 and 1.2 mg/dL and was lower in those with serum creatinine values less than 0.9 or greater than 1.2 mg/dL (Appendix Figure 2). Similar relations were seen between survival and heart rate, serum sodium level, temperature, and respiratory rate.

The relation between survival and some variables varied with disease group. For example, in patients with lung or colon cancer, lower serum albumin levels had an especially strong relation with risk for death. A similar but weaker association was seen in patients with chronic obstructive pulmonary disease and congestive heart failure, but serum albumin values were not associated with risk in the other diagnostic groups. Another significant disease-specific interaction was low leukocyte count, which was associated with a greater risk for death in patients with acute respiratory failure and those with multiple organ system failure (Appendix Figure 2).

The Glasgow coma scale was the single physiologic risk factor most predictive of risk for death (Figure 1); it had a nonlinear relation to this risk. Using scores derived from independent database APACHE III (4), we rescaled the Glasgow coma scale so that it could have a nonlinear relation to risk for death. In the revised scale, 0 is equivalent to a Glasgow coma scale value of 15 (normal), 100 is equivalent to a Glasgow coma scale value of 3 (deep coma), and 44 is equivalent to a Glasgow coma scale value of 9 (intermediate) (Appendix Figure 2). Patients for whom a Glasgow coma scale value could not be reliably assessed were assigned a normal Glasgow coma scale value.

The SUPPORT physiology score consists of 11 physiologic variables, each with a continuous weighting scheme (Appendix Figure 2). With the exception of the Glasgow coma scale component and some disease-specific scorings, the relative shapes of the weightings assigned to the physiologic variables in the SUPPORT physiology score were similar to those used in the APACHE III acute physiology score (4); a high correlation existed between the two scores ($r = 0.78$).

Long-Term Health Evaluation and Previous Hospital Days

The overall effect of age on survival in our study was modest and similar to that in the APACHE III system (4). However, we found a significant interaction when we allowed the influence of age to vary for the nine disease groups. The age effect was greatest in patients with chronic obstructive pulmonary disease, in whom an increase in age from 70 to 75 years increased the risk for death within 180 days by approximately 10%. Risk for death among patients with multiple organ system failure and malignancy was not substantially affected by age (Appendix Figure 3). Because of our study entry criteria, the only comorbidity that was a statistically significant risk factor was diagnosis of any malignancy. The SUPPORT prognostic equation also includes a variable for the number of days the patient was hospitalized before study entry (Appendix Figure 4).

The final SUPPORT prognostic model (Appendix Figures 1–4) contains the three disease-specific interactions described previously and 15 prognostic factors: disease group, 11 physiologic variables, patient age, history of malignancy, and the number of days the patient was hospitalized before study entry. Full equations for deriving the SUPPORT physiology score and the day 3 prognostic model appear in Appendix 2.

Model Validation Using Phase II Data

Independent validation of the model's performance on the 4028 phase II patients who had 180-day follow-up and physician's independent estimates showed good calibration and a ROC curve area of 0.78 (Table 2; Figure 2, *top*). The model's performance did not vary significantly across the five sites in either the phase I or the phase II data sets.

Comparison of SUPPORT and APACHE III Model Predictions

Comparison of the revised weighting of the physiologic variables in the SUPPORT prognostic model with the APACHE III prognostic model was done by substituting the APACHE III acute physiology score for the 11 physiologic variables included in the SUPPORT physiology score (Appendix Figure 5, *bottom*) in phase I, and refitting the Cox model. In this comparison in phase II pa-

Table 2. Comparison of the Various Models for Prediction of 180-Day Survival*

Disease class	SUPPORT Model	SUPPORT Model with APS†	Physician's Estimate	SUPPORT Model and Physician's Estimate
All (n = 4028, deaths = 1899)	0.78	0.78	0.78	0.82
Acute respiratory failure and multiple organ system failure (n = 2057, deaths = 993)	0.77	0.78	0.78	0.82
[illegible] obstructive pulmonary disease congestive [illegible] failure, cirrhosis (n = 1111, deaths = 346)	0.71	0.70	0.70	0.75
Coma (n = 281, deaths = 205)	0.74	0.75	0.78	0.82
Colon and lung cancer (n = 579, deaths = 345)	0.78	0.70	0.77	0.82

* All calculations are based on 4028 SUPPORT phase II patients who completed 180 days of follow-up and had a physicians' prognostic estimate at study day 3. Each statistic is the area under the receiver-operating characteristic curve for 180-day vital status.
† APS = APACHE III acute physiology score.

tients, the ROC curve area of 0.78 for all patients did not change (Table 2; Figure 2, *middle*), but for certain types of patients, especially those with chronic conditions such as cancer, the SUPPORT model showed important improvements, evidenced by the larger ROC curve area (Table 2).

Comparison of SUPPORT Model and Physicians' Predictions

For the 4028 phase II patients with both SUPPORT model and physician estimates, the ROC curve area for both types of estimate were identical, 0.78 (Table 2). The physicians were slightly more pessimistic than the SUPPORT model and also gave more extreme estimates of survival or death when estimating very low probabilities of either outcome (Figure 2, *bottom*).

When the physicians' subjective estimates were incorporated into the SUPPORT model as an additional variable, significant explanatory power was added to the model's estimate of 180-day survival (ROC curve area = 0.82). Table 2 shows that the highest ROC curve area is achieved when the SUPPORT model is combined with the physicians' estimates of 180-day survival.

The advantage of combining objective prognostic estimates with physicians' clinical judgments is further shown in Figure 3, which compares the predictive ability of the SUPPORT prognostic model, with and without substitution of the APACHE III acute physiology score, the physician's prediction, and the physician-enhanced SUPPORT model for the 4028 phase II patients. Prognostic discrimination was compared in the patient samples predicted to have a low likelihood of survival (survival probabilities lower than 0.15; Figure 3, *top*) and to be very likely to survive (greater than 0.85; Figure 3, *bottom*) by the four prediction methods.

Physicians identified more patients at high risk for dying (<0.15 likelihood of survival) than did the SUPPORT model (n = 753 for physicians; n = 471 for the SUPPORT model), with an observed mortality rate of 15% for physicians and 12% for the SUPPORT model. In comparison, the physician-enhanced SUPPORT model identified 668 patients at high risk for dying, and the observed survival rate of 11% was the lowest in that risk interval among the three models (Figure 3, *top*). Among patients very likely to survive (likelihood of survival >0.85), the physician-enhanced SUPPORT model identified an equivalent number of patients (n = 496; n = 484 for physicians) with slightly greater accuracy than did physicians alone (actual 180-day survival = 91% compared with 85% for physicians; Figure 3, *bottom*).

A similar complementary pattern appeared when we compared the incremental value of the SUPPORT model with physicians' estimates. Both contained substantial independent prognostic information for all patients and for all four disease classes: The SUPPORT prognostic model added a chi-square of 519 to the physicians' estimates and the physicians added a chi-square of 522 to the SUPPORT prognostic model ($P < 0.001$).

Within the independent phase II validation, the SUPPORT prognostic model, the SUPPORT prognostic model with the acute physiology score inserted, and the physicians' subjective estimates (Figure 2) were all overly pessimistic when predicting low probabilities of survival. In response, we investigated 27 patients who survived 180 days and for whom the SUPPORT model had estimated a 180-day survival probability of less than 0.05. This investigation showed that the low prognostic estimate was due either to 4 patients who had a cardiac arrest on day 3, 3 patients who subsequently received a liver transplant, or 4 patients who died shortly after the 180-day follow-up. These results emphasize the need for caution when using prognostic estimates based on physiology during a cardiac arrest and the ability of new, highly efficacious therapies, such as organ transplantation, to improve prognosis. They also emphasize the limitations of a prognostic estimate done on a single day, as opposed to estimates done over time, and the need to always integrate the results of any clinical measurement within the overall clinical context.

Discussion

Reasoning based on event probabilities has been introduced into many scientific disciplines (19). In medicine, accurate probabilities are available for diagnostic challenges, for example, the probability that a patient with chest pain will have a myocardial infarction (20, 21) and the likelihood of death from coronary artery bypass surgery (22, 23). Clinicians, however, do not now incorporate such prognostic data into routine decision making; many are concerned that a few selected characteristics drawn from a group of previously treated patients might not accurately reflect an individual patient's risk (24). Prognostic estimates are also not yet generally available at the time of decision making. Because the technical ability to

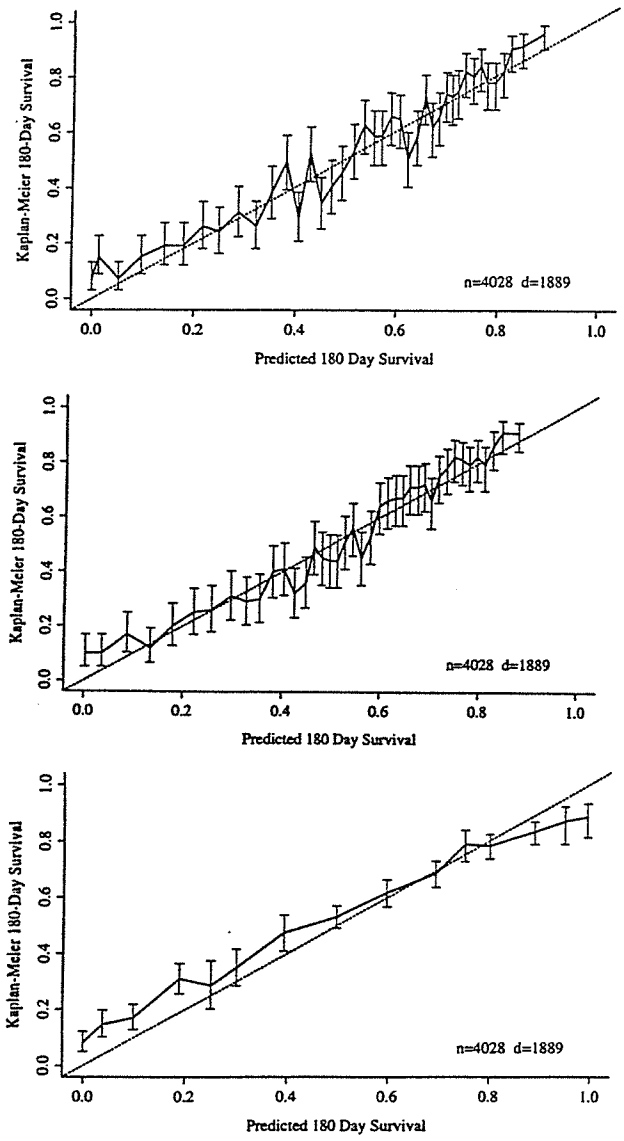


Figure 2. Prospective validation of calibration of the various prognostic models based on day 3 data for survival to 180 days after study entry for 4028 Phase II patients. The fraction surviving is on the vertical axis and the model prediction is on the horizontal. Each bar represents the mean prediction for 100 patients in each interval; the heights of the bars indicate 95% CIs. The overall calibration or reliability of the model is expressed as the closeness of the fit of this curve to the diagonal, which is the ideal fit. **Top. SUPPORT prognostic model.** Receiver-operating characteristics curve area for 180-day survival = 0.78. **Middle. SUPPORT prognostic model with the APACHE III acute physiology score substituted for the SUPPORT physiology score.** Receiver-operating characteristics curve area for 180-day survival = 0.78. **Bottom. Physicians' estimates.** There are fewer bars because physicians did not use all possible probabilities in forming predictions; they tended to use multiples of 0.05. Receiver-operating characteristics curve area = 0.78. Average number of predictions for physicians = 269.

provide such estimates is improving rapidly (24, 25), and because many decisions for seriously ill hospitalized adults are based in part on the risk for death, a more explicit and informed statement of these risks may be helpful (26). Most adults in the United States say that if their deaths could occur in less than a year, they want a realistic estimate of how long they can expect to live (27). Avoiding discussions of prognosis can make patients feel abandoned and physicians feel estranged (28, 29). Additionally, the evolving professional and societal consensus is that prognostic data should be shared in the context of the patient–physician relationship, which is based on trust (30–32).

The results of our study indicate that within a selected group of high-mortality diagnoses, a limited amount of readily available clinical information can produce probabilities of survival that are as accurate as those by treating physicians (Table 2, Figure 2). Our results also suggest that these objective probability estimates may complement physicians' prognoses. The improvement in accuracy at low probabilities of survival may be helpful in refining decisions to limit or withdraw life support (33), but the previously mentioned limitations of relying on a single prognostic estimate deserve emphasis (25). The enhanced discrimination gained when physician and model estimates are combined (Figure 3), however, is especially noteworthy because it could increase physicians' confidence in prediction at both extremes of risk and could thereby avoid both unnecessarily optimistic and pessimistic prognoses. The acceptance and understanding of these estimates by clinicians was initially evaluated during phase II of SUPPORT. An example of this feedback report for a patient with multiple organ system failure with malignancy is provided in Appendix Figure 5, *top*. This prognostic report provided both the estimate and the contribution of each individual prognostic factor to that estimate (Appendix Figure 5, *bottom*).

Prognostic Factors and the SUPPORT Model

In our model, the single most important prognostic factor was severity of the physiologic responses (Figure 1). The total chi-square attributable to these physiologic abnormalities was 1058 compared with 168 for specific disease groups, confirming previous reports (4–6, 34) and emphasizing the need to take acute physiologic response into account when assessing seriously ill hospitalized patients (19, 28). Our results also suggest that in selected disease groups, some physiologic factors may be more important to prognosis than others. A low serum albumin level, for example, was of substantial prognostic importance in patients with metastatic colon or lung cancer and was important in those with congestive heart failure and chronic obstructive pulmonary disease, but it was relatively unimportant in patients with coma, respiratory failure, or multiple organ system failure (Appendix Figure 2). One possible explanation for this is that serum albumin values represent physiologic reserve (for example, nutritional status in patients with cancer, congestive heart failure, or other chronic conditions) and represent other phenomena, such as rapid fluid shifts, in more acute processes, such as acute respiratory failure or multiple organ system failure (35). Our results confirm the long-recognized relation between a low leukocyte count and a poor prognosis for recovery in patients with severe respiratory failure (36), and they extend this relation to include patients with other acute organ system failures. This is consistent with recent evidence that a low leukocyte count has important prognostic implications for patients with the sepsis syndrome (37).

We examined many established disease-specific prognostic variables, such as ejection fraction for patients with congestive heart failure, expecting that they might enhance prognostic estimates derived from general prognostic variables alone (38). They did not, probably because these measurements were frequently missing at the time prognostic estimates were generated. In addition, our selection criteria may have identified a cohort of patients whose illness was so severe that the influence of these variables was less important, and the prognostic effect of the disease-specific factors may have been captured by the generic physiologic measures. A more accurate test of the predictive value of disease-specific measurements would require a more uniform collection of these measurements than was possible in this observational study.

Previous studies have shown the other nonphysiologic elements of the SUPPORT model—chronologic age, presence of cancer, and number of days in the hospital before study entry—to have independent prognostic value (4–6, 34, 39). Their relative value is similar to that of physiologic and disease variables in the SUPPORT model. However, we permitted the influence of some of these factors, such as age, to vary depending on the disease class, acknowledging that although increases in chronologic age result in decreases in physiologic reserve, the incremental influence on the short-term risk for death is small. In serious acute complications of a chronic illness, such as multiple organ system failure in patients with cancer, age has little prognostic significance because acute physiologic severity determines outcome (Appendix Figure 3). These findings may influence policy by assuring that chronologic age is used appropriately in treatment guidelines and medical decisions (24).

To extract the maximum information value from each of the prognostic factors in this study, we used a relatively new statistical technique, restricted cubic splines, to produce a continuous weighting scheme. This approach makes few a priori assumptions about the shape of the risk relation between physiology and outcome (Appendix Figure 2) and allows a flexible and biologically appropriate representation of the relation between prognostic variables and risk compared with the point scores used, for example, in the APACHE III system (4). This method uses computers to derive all variable weightings in the model and to calculate the prognostic estimate. The emphasis on computerization and the model's reliance on advanced statistical techniques may raise concern that this approach is not as useful as simpler prediction rules (40). Compared with simpler approaches, however, this method can combine the complex manifestations of illness by using a wide range of prognostic components, from serum albumin levels to chronologic age, and can take into account their influence within specific disease classes. This method can provide a systematic assessment of a patient's prognosis and assist in educating the physician about the relative importance of each prognostic element (Appendix Figure 5, *bottom*). Decision rules or prognostic systems that rely on only a few selected variables make the assumption that variables not included are not influential in prognosis. Clinical decision making, however, attempts to integrate all available data, especially established prognostic factors (41). Although the incremental importance of a larger number of variables or a more complex mathematical treatment may not, as was the case here, increase overall explanatory power for all patients combined (ROC curve area of 0.78 when using the SUPPORT physiology score compared with 0.78 when using the APACHE III acute physiology score; Table 2), it may make a difference in predictions for many individual patients, including those at both extremes of the risk profile (Figure 3) or within specific disease groups (Table 2).

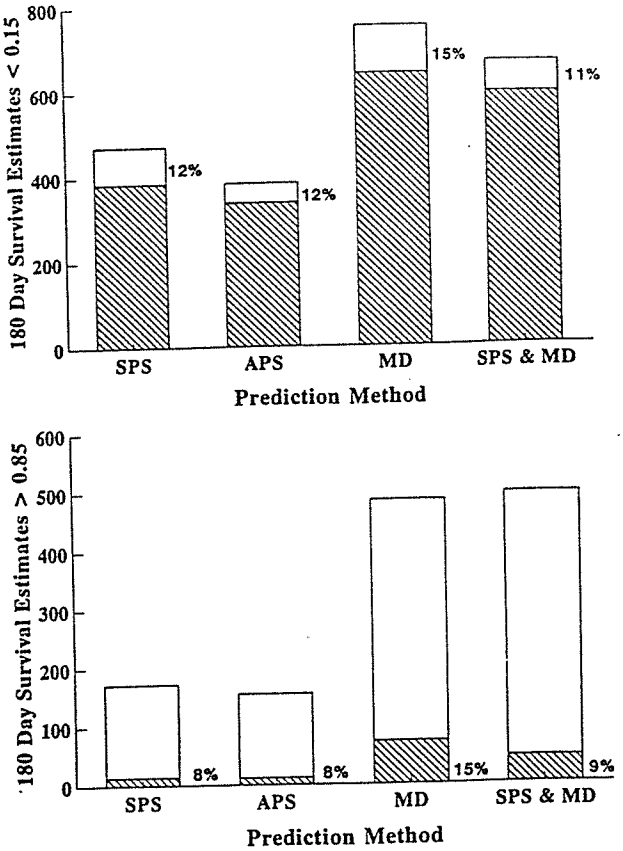


Figure 3. Comparison, in high and low ends of the prognostic range, of the predictive ability of four prognostic models. SPS = SUPPORT prognostic model; APS = SUPPORT prognostic model with the APACHE III acute physiology score; MD = physician's prediction; and SPS & MD = physician-enhanced SUPPORT model. Results are based on 4028 phase II patients who were followed for 180 days and for whom physician estimates were obtained. The vertical axes represent the number of patients in each sample. Cross-hatched sections indicate the proportion of patients who died; open sections indicate survivors. **Top. Number of phase II patients with a 180-day survival probability of less than 0.15.** Physicians predicted a survival less than 0.15 for more patients (n = 753), but within this risk group, had the highest survival rate and the lowest calibration of the three prognostic models. The SUPPORT physician-enhanced model had the lowest survival rate when predicting low probability of survival and the second highest number of patients (n = 668). **Bottom. Number of phase II patients with a survival probability of greater than 0.85.** The physician-enhanced prognostic model predicted the largest sample (n = 496) and had a 91% survival rate; the physician-estimated sample had an 85% mortality rate.

Individual Estimates Compared with Prediction Rules

One important reason to consider introducing numeric estimates of a patient's individual probability of outcome or survival time is that decisions likely to be influenced by

this estimate, such as those about the appropriateness of continued or new therapy, are directly related to the overall burden or severity of illness. A prediction rule that groups patients together may group persons with a wide range of risks (42); this may not be helpful in making difficult decisions involving trade-offs between risks and benefits that are related to specific levels of risk. We also note that, despite their simplicity, few of the many proposed clinical prediction rules are actually used at the bedside. The reluctance to use them may be due to concerns that they do not reflect state-of-the-art therapeutic capabilities (43). Computerized prognostic estimates derived from large contemporary clinical databases similar to the one described here can provide estimates as quickly as laboratory tests can, usually within a few minutes after data collection. The databases that support these estimates can be constantly updated, ensuring that the prognostic estimate is compatible with current therapy. Indeed, such estimates may be useful in determining the incremental value of new therapies (37).

Limitations of the SUPPORT Model and Future Development

The selection criteria for any prognostic system will always have a direct and important effect on the model's performance and subsequent validation. Although the SUPPORT prognostic model applied to only a small percentage of admissions in the five study hospitals, these patients accounted for approximately one fifth of all in-hospital deaths. Because this model was developed at only a few teaching hospitals, on a sample of patients who met select entry criteria, its usefulness at other hospitals awaits validation. If the level of performance shown in this study is maintained when the model is applied to a less restrictive database with more easily replicated selection criteria, the method may be a powerful foundation for making decisions about individual patients. The current SUPPORT prognostic model, however, has been shown to be accurate only with patients who are selected under our specific entry criteria (Appendix 1). This model is also restricted to data available by the end of the third day after qualifying for the study, although complementary models for other days are available (Appendix 2). These restrictions may make it difficult to implement this approach on a routine basis; if this is done, local performance should be carefully monitored. Methods used in its development, however, should prove useful for the development of future prognostic systems on larger, more representative databases. The construction of the model also provides flexibility, enabling it to incorporate new disease-specific prognostic findings as they become available (44).

A particularly promising finding of our study is that the SUPPORT model complements simultaneous physicians' prognostic estimates (Figure 3). Previous efforts to estimate risk for death have been criticized because the objective estimates were not superior to clinical judgment (45–47). When designing the SUPPORT study, however, we presumed that the prognostic estimates would be used in conjunction with—and not as a replacement for—physicians' estimates. The investigation of the relative strengths and weaknesses of physicians' prognostic estimates, and of how to optimize the model's contribution to physicians' estimates, will require extensive additional investigation (26). There are many theoretical advantages to physicians' estimates. Besides being familiar with elements of the patient's condition not included in the model, physicians may be able to provide superior imputations of missing data for an individual patient and to integrate the risk estimate as part of their overall patient assessment. In our statistical analysis, the SUPPORT model used more of the prognostic range than did physician estimates, which tended to cluster certain probabilities and made slightly more predictive errors at the extremes of risk. When both the SUPPORT model and the physician indicated either a high or a low probability of survival, however, predictive accuracy exceeded that of clinical judgment alone (Figure 3).

The SUPPORT prognostic model provides a new, accurate, and flexible empiric method for estimating a patient's risk for death over time. Its results equal physicians' estimates using easily obtained and reproducible clinical measures.

Appendix 1: Inclusion and Exclusion Criteria for the Nine Disease Groups Used in SUPPORT

Patients were placed in the first category (the one with the lowest number) for which they qualified. They could meet the necessary criteria at time of hospitalization or at any time during treatment in an intensive care unit.

1. Nontraumatic coma

Inclusion criteria: Documentation of "coma" or "unresponsive" defined as a Glasgow coma scale score of 9 or less, lasting for 6 or more hours.

Exclusion criteria: Evidence of drug intoxication, hypothermia, or metabolic disturbances (except hypoglycemia and hypoxemia) as the primary cause of coma, general anesthesia within the previous 48 hours, determination of brain death within 48 hours of onset of coma, or patients who had a normal preoperative neurologic examination but who remain unresponsive after surgery.

2. Multiple organ system failure and malignancy

Inclusion criteria: Care in the intensive care unit, an APACHE II acute physiology score of 15 or more (12 or more if paralyzed with medications) at admission, and documentation of any solid or hematologic malignancy currently present in at least one site distant to the original location.

Exclusion Criteria: Multiple traumas, near-drowning, drug intoxication, or primary hypoventilation including that associated with the Guillain-Barré syndrome.

3. Acute respiratory failure

Inclusion criteria: Admission to the intensive care unit, documentation of suspected pneumonia or the adult respiratory distress syndrome, and an APACHE II acute physiology score of 10 (7 or more if paralyzed with medications).

Exclusion criteria: Severe chronic obstructive pulmonary

disease or congestive heart failure, *Pneumocystis carinii* pneumonia, status asthmaticus, pulmonary embolism, immunologic lung disease, primary restrictive lung disease, primary hypoventilation including that associated with the Guillain-Barré syndrome, smoke inhalation, or a thoracotomy during current hospitalization.

. Multiple organ system failure and sepsis

Inclusion criteria: Care in the intensive care unit, an APACHE II acute physiology score of 15 or more (12 or more if paralyzed with medications) at admission, and a clinical impression of sepsis, septicemia, or bacteremia.

Exclusion criteria: Near-drowning or drug intoxication.

5. Acute exacerbation of severe chronic obstructive pulmonary disease

Inclusion criteria: Clinical diagnosis of chronic obstructive pulmonary disease, chronic bronchitis, chronic obstructive lung disease, or emphysema with breathlessness, respiratory failure, or mental status change as the main reason for hospital admission, and hypercapnea and hypoxemia ($Po_2 \leq 60$ mm Hg and $Pco_2 \geq 50$ mm Hg if the patient is receiving room air, or $Pco_2 \geq 50$ mm Hg alone if the patient is receiving supplemental oxygen) documented at admission.

Exclusion criteria: Status asthmaticus.

6. Acute exacerbation of severe congestive heart failure

Inclusion criteria: Clinical diagnosis of congestive heart failure or cardiomyopathy with an exacerbation of symptoms as the primary reason for hospital admission plus one of the following: 1) a history of severe congestive heart failure at baseline (New York Heart Association class III or IV) manifested by a history of dyspnea at rest or with minimal exertion related to primary cardiac failure, and medications before admission that include at least two drug classes (diuretics, vasodilators, or adrenocortical extract inhibitors), and a history of class III or IV congestive heart failure at hospital admission documented by dyspnea at rest at baseline; 2) a history of class IV congestive heart failure at admission, dyspnea at rest, and systolic blood pressure of 100 mm Hg or less, or a history of hypotension that precludes the use of diuretics, vasodilators, or adrenocortical extract inhibitors; and 3) documentation of severe congestive heart failure with an ejection fraction of 20% or less.

Exclusion criteria: Severe chronic obstructive pulmonary disease, shock, primary acute renal failure, decreased systemic vascular resistance, restrictive cardiac disease, circulatory overload, congestive heart failure resulting primarily from valvular heart disease, cardiac surgery, or thoracotomy during current hospitalization.

7. Chronic liver disease

Inclusion criteria: Chart documentation of cirrhosis and at least two of the following: a serum albumin level of 3.0 mg/dL or less, a serum bilirubin level of 3.0 mg/dL or more, uncontrolled ascites, hepatic encephalopathy, documentation of cachexia, or a massive gastrointestinal bleed defined as two or more blood transfusions in 24 hours and either hematemesis or gross blood on endoscopic visualization or nasogastric tube aspiration.

8. Colon cancer with liver metastasis

Inclusion criteria: Known cancer of the colon or rectum and metastasis to the liver at hospital admission.

Exclusion criteria: New diagnosis within the previous 30 days and first hospitalization for cancer.

9. Non–small cell carcinoma of the lung

Inclusion criteria: Documentation of non–small cell carcinoma of the lung at hospital admission, and stage III or IV disease manifested by known involvement of the mediastinum, hilum, or peribronchial nodes, or known involvement of the pleural space, or known distant metastases.

Exclusion criteria: Cell types other than squamous cell or adenocarcinoma of the lung, new diagnosis within the previous 30 days and first hospitalization for cancer.

Appendix 2*

Formula for computing the SUPPORT physiology score and the entire SUPPORT day 3 prognostic model.

SUPPORT Physiology Score (Range, 0–100):

$$\begin{aligned}SPS = {} & 259.9\{ARF/MOSF\} + 263.4\{COPD/CHF\} + 241.4\{Cirrhosis/Coma\} + 281.5\{Lung/Colon\ Cancer\} \\ & - 0.06174\ \min(PaO_2/FiO_2, 225) - 0.6316\ \min(Mean\ BP, 60) + 1.0205\ WBC \\ & - 0.3676(WBC - 8)_+ - 0.5631(WBC - 11)_+ + 0.2691\ \min(Alb, 4.6) + 0.2312\ Aresp \\ & - 2.362\ Temp + 1.326(Temp - 36.6)_+ + 2.473(Temp - 38.3)_+ \\ & - 1.579 \times 10^{-1}\ HR + 9.770 \times 10^{-5}\ (HR-55)^3_+ - 2.189 \times 10^{-4}(HR - 80)^3_+ \\ & + 1.518 \times 10^{-4}(HR - 110)^3_+ - 3.062 \times 10^{-5}(HR - 149)^3_+ + 0.9763\ Bil \\ & - 0.7481(Bil - 7)_+ - 6.8761\ Cr + 11.6058(Cr - 0.600)^3_+ - 21.8413(Cr - 1.000)^3_+ \\ & + 10.3574(Cr - 1.500)^3_+ - 0.1219(Cr - 5.399)^3_+ - 0.6167096\ Na \\ & + 0.0021118(Na - 128)^3_+ - 0.0036730(Na - 135)^3_+ + 0.0006126(Na - 139)^3_+ \\ & + 0.0009486(Na - 148)^3_+ - 6.278\ \{COPD/CHF\} \times \min(Alb, 4.6) \\ & - 11.45\ \{Lung/Colon\ Cancer\} \times \min(Alb, 4.6) + \{ARF/MOSF\}[-2.3549\ WBC \\ & + 2.7494\ (WBC - 8)_+ - 0.4638\ (WBC - 11)_+]\end{aligned}$$

where {disease group} = 1 if subject is in the disease group, 0 otherwise, $(x)_+ = x$ if $x > 0$, 0 otherwise. For example, the term $0.4638\ (WBC - 11)_+$ is ignored if $WBC \leq 11$. These terms are components of cubic spline functions. All measurements are made at day 3 except albumin (day 1). Alb: albumin; Aresp: APACHE III respiration score; Bil: bilirubin; Cr: creatinine; Na: sodium; P_aO_2: partial pressure oxygen in arterial blood; Mean BP: mean arterial blood pressure; WBC: white blood cell count in thousands; Temp: temperature (Celsius); HR: heart rate per minute. Set WBC = 9 in WBC < 9 and the disease class is not ARF/MOSF. Set WBC = 40 if WBC > 40. Set Cr = 15 if Cr > 15.

$$\text{Probability}\{T \geq t \text{ given disease class} = i\} = S_i(t)^{e^{x\beta}},$$

where

T = survival time in days, t = an arbitrary time, e is the

SUPPORT Day 3 Prognostic Model

t	$S_{ARF/MOSF}(t)$	$S_{COPD/CHF/Cirrhosis}(t)$	$S_{Coma}(t)$	$S_{Cancer}(t)$
0	0.994	0.998	0.993	0.993
30	0.691	0.889	0.630	0.578
60	0.601	0.837	0.609	0.407
90	0.562	0.800	0.581	0.264
120	0.532	0.772	0.569	0.190
150	0.508	0.751	0.551	0.135
177	0.493	0.733	0.545	0.108

base of the natural logarithm, and $X_{\hat{\beta}} = -3.652 + 0.8353$ {CHF} + 0.9257 {Cirrhosis} + 0.6287 {Lung Cancer} ± 1.1803 {MOSF w/Malig} + 0.01434 Scoma ± 0.01935 Age + 0.2413 Cancer − 1.863 [Hday + 3.4]$^{-1}$ + 0.08121 SPS + Age[0.015261 {COPD/CHF/Cirrhosis} + 0.009047 {Coma} − 0.008294 {Cancer}] + Age[−0.012498 {CHF} − 0.004578 {Cirrhosis} − 0.001435 {Lung Cancer} − 0.013891 {MOSF w/Malig}] and {disease group} = 1 if subject is in the disease group, 0 otherwise.

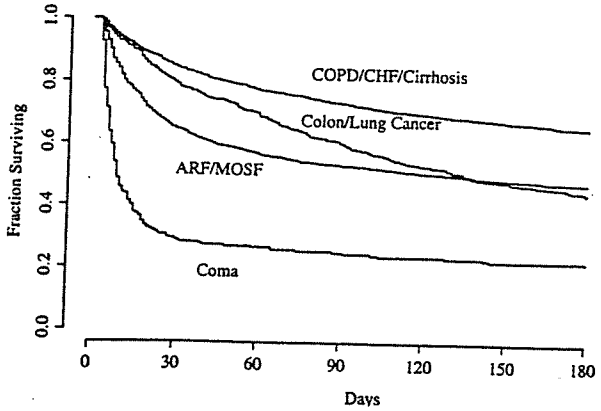
Scoma = SUPPORT Coma Score (0–100); cancer = cancer by comorbidity or primary disease category (0 = no; 1 = present; 2 = metastatic); Hday = day in hospital when qualified for study; CHF = congestive heart failure; MOSF w/Malig = multiple organ system failure with malignancy; COPD = chronic obstructive pulmonary disease.

*Equation:

Copies of the SUPPORT prognostic models for days 1, 7, 14, and 25 are available from the authors.

Grant Support: By The Robert Wood Johnson Foundation. The opinions and findings contained in this article are those of the authors and do not necessarily represent the views of The Robert Wood Johnson Foundation or their Board of Trustees.

Requests for Reprints: William A. Knaus, MD, National Coordinating Center, The Robert Wood Johnson Foundation Critically Ill Hospitalized Adult Program, ICU Research Unit, 2300 K Street NW, Washington, DC 20037.



Appendix Figure 1. Relation between disease classification in the SUPPORT prognostic model and proportion of patients surviving to 6 months. The 6-month mortality in the 4301 phase I SUPPORT patients was 48.1%, but because the shapes of survival curves varied substantially, the nine disease groups (Appendix 1) were collapsed into four classes. ARF = acute respiratory failure, CHF = congestive heart failure, COPD = chronic obstructive pulmonary disease, MOSF = multiple organ system failure.

Current Author Addresses: Dr. Knaus: National Coordinating Center, The Robert Wood Johnson Foundation Critically Ill Hospitalized Adult Program, ICU Research Unit, 2300 K Street NW, Washington, DC 20037.

Dr. Harrell: Division of Biometry and the Heart Center, Duke University Medical Center, Erwin Square, Bay B, Room 215, 2024 West Main Street, Durham, NC 27705.

Dr. Lynn: Dartmouth-Hitchcock Medical Center, Center for Evaluative Clinical Sciences, Dartmouth Medical School, 7250 Strasenburgh Hall, Room 312, Hanover, NH 03755-3863.

Dr. Goldman: Brigham and Women's Hospital, Old PBB3 Building Executive Offices, 75 Francis Street, Boston, MA 02215.

Dr. Phillips: Beth Israel Hospital, LY-330, 330 Brookline Avenue, Boston, MA 02215.

Dr. Connors and Dawson: Cleveland Metrohealth Medical Center, 2500 Metrohealth Drive, Rammel Kamp Center, Cleveland, OH 44109-1998.

Dr. Fulkerson: Duke South Hospital, Room 0522, Blue Zone, Trenton Drive, Duke University Medical Center, Durham, NC 27710.

Dr. Califf: Cardiac Data Bank, Duke University Medical Center, Erwin Square, Bay A108, 2024 West Main Street, Durham, NC 27710.

Dr. Desbiens: Marshfield Medical Research Foundation/Marshfield Clinic, 1000 North Oak Avenue -1R5, Marshfield, WI 54449-5790.

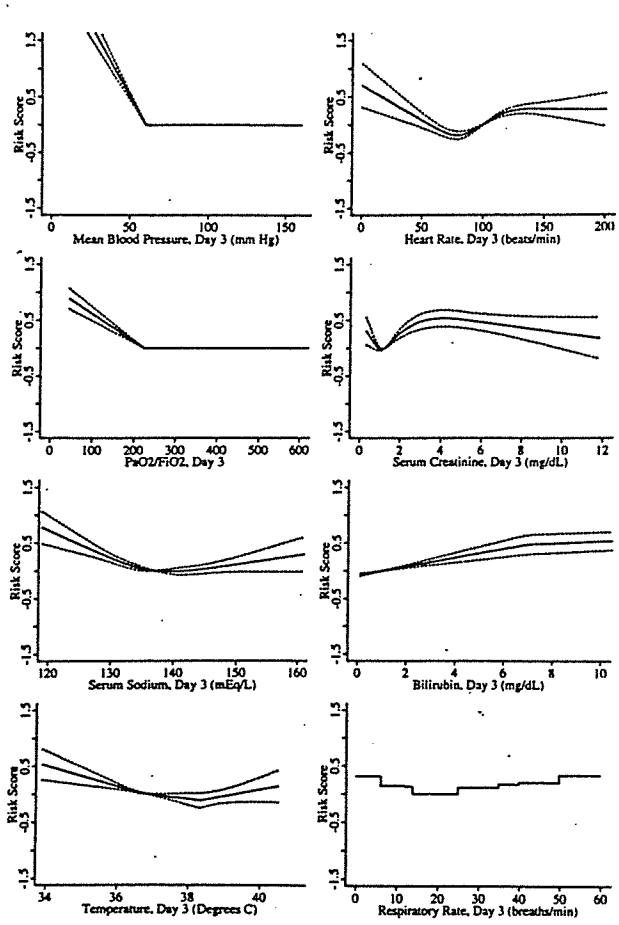
Dr. Layde: Medical College of Wisconsin, Department of Family and Community Medicine, 8701 Watertown Plank Road, Milwaukee, WI 53226.

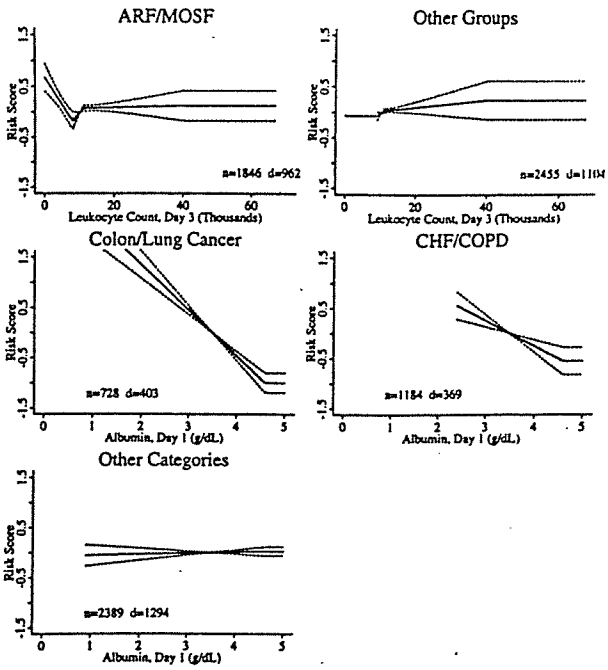
Drs. Oye and Bellamy: Department of Medicine, University of California at Los Angeles School of Medicine, 10833 Le Conte Avenue, Los Angeles, CA 90024-1678.

Drs. Hakim and Wagner: National Coordinating Center, ICU Research Unit, Warwick Building, 2300 K Street NW, Washington, DC 20037.

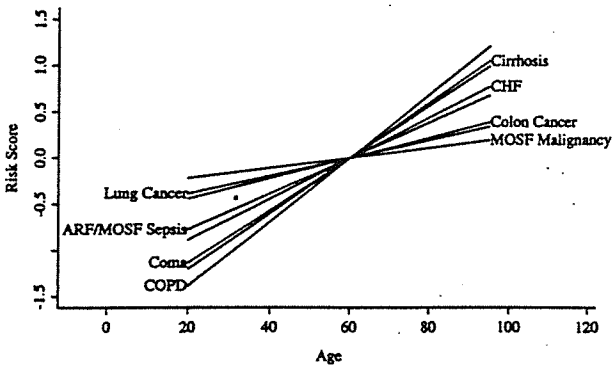
References

1. **Murphy DJ, Cluff LE.** The SUPPORT study. Introduction. J Clin Epidemiol. 1990;43(Suppl):V-X.
2. **Connors AF Jr, Dawson NV, Lynn J.** Decision making in SUPPORT: introduction. J Clin Epidemiol. 1990;43(Suppl):47S-9S.
3. **Lynn J, Knaus WA.** Background for SUPPORT. J Clin Epidemiol. 1990;43(Suppl):1S-4S.
4. **Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al.** The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest. 1991; 100:1619-36.
5. **LeGall JR, Lemeshow S, Saulnier F.** A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. JAMA. 1993;270:2957-63.
6. **Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J.** Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. JAMA. 1993;270:2478-86.
7. **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.
8. **Murphy DJ, Knaus WA, Lynn J.** Study population in SUPPORT: Patients (as defined by disease categories and mortality projections), surrogates, and physicians. J Clin Epidemiol. 1990;43(Suppl):11S-28S.
9. **Cox DR.** Regression models and life tables (with discussion). J R Stat Soc B. 1972;34:187-220.
10. **Harrell FE, Lee KL.** Verifying assumptions of the Cox proportional hazards model. In: Proceedings of the Eleventh Annual SAS Users Group International Conference (Cary, North Carolina), 1986; 823-828. SAS Institute, Inc.
11. **Schoenfeld D.** Partial residuals for the proportional hazards regression model. Biometrika. 1982;69:239-241.
12. **Harrell FE Jr, Lee KL, Pollock BG.** Regression models in clinical studies: determining relationships between predictors and response. J Natl Cancer Inst. 1988;80:1198-202.
13. **Durrleman S, Simon R.** Flexible regression models with cubic splines. Stat Med. 1989;8:551-61.
14. **Stone CJ, Koo CY.** Additive splines in statistics. In: Proceedings of the Statistical Computing Section of the American Statistical Association. 1985; 45-48.
15. **Van Houwelingen JC, Le Cessie S.** Predictive value of statistical models. Stat Med. 1990;9:1303-25.
16. **Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA.** Evaluating the yield of medical tests. JAMA. 1982;247:2543-6.
17. **Harrell FE, Lee KL.** A comparison of the discrimination of discriminant analysis and logistic regression under multivariate normality. In Sen PK, ed. Biostatistics: Statistics in Biomedical, Public Health, and Environmental Sciences. North-Holland, New York: Elsevier; 1985: 333-43.
18. **Damiano AM, Bergner M, Draper EA, Knaus WA, Wagner DP.** Reliability of a measure of severity of illness: acute physiology and chronic health evaluation—II. J Clin Epidemiol. 1992;45:93-101.
19. **Dawes RM.** Rational Choice in an Uncertain World. New York: Harcourt Brace & Jovanovich; 1988.
20. **Pozen MW, D'Agostino RB, Selker HP, Sytkowski PA, Hood WB Jr.** A predictive instrument to improve coronary-care-unit admission prac-

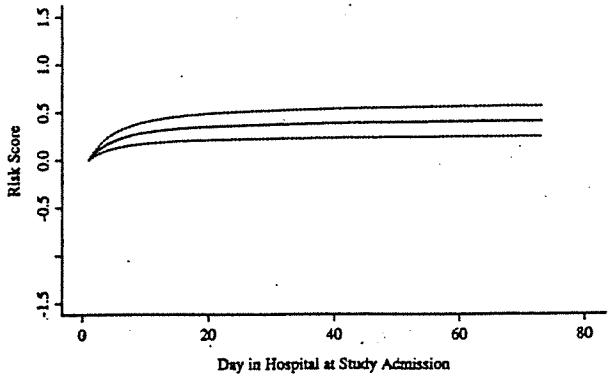




Appendix Figure 2. SUPPORT Physiology Score in the SUPPORT prognostic model. The SUPPORT physiology score consists of 11 physiologic measures taken on study day 3, represented on a continuous basis. The risk score is equivalent to the log relative hazard. A risk score of 0 indicates that the value of the variable places the patient at the "standard" or baseline risk, and a positive or negative score indicates an increased or decreased risk, respectively, relative to baseline. Dotted lines indicate 95% CIs. The relative risk assigned to respiratory rate was adapted directly from the APACHE III prognostic scoring system, modified by a regression coefficient estimated in SUPPORT. n = the number of patients in the subgroup; d = the number of patients in the subgroup who had died by 180 days.



Appendix Figure 3. Long-term health evaluation in the SUPPORT prognostic model. The two components of this factor are chronologic age and the presence of cancer as a comorbidity; the interaction between age and disease is illustrated. Disease groups were combined according to the relative effect of age into multiple organ system failure with cancer; lung and colon cancer; acute respiratory failure (*ARF*) or multiple organ system failure (*MOSF*) without cancer, and cirrhosis and coma; congestive heart failure (*CHF*); and chronic obstructive pulmonary disease (*COPD*). Mean patient age was 63 years. Because the selection criteria required patients to have severe primary disease (Appendix 1), the only comorbidity that had independent prognostic significance was cancer; metastatic solid tumors had the highest relative risk.



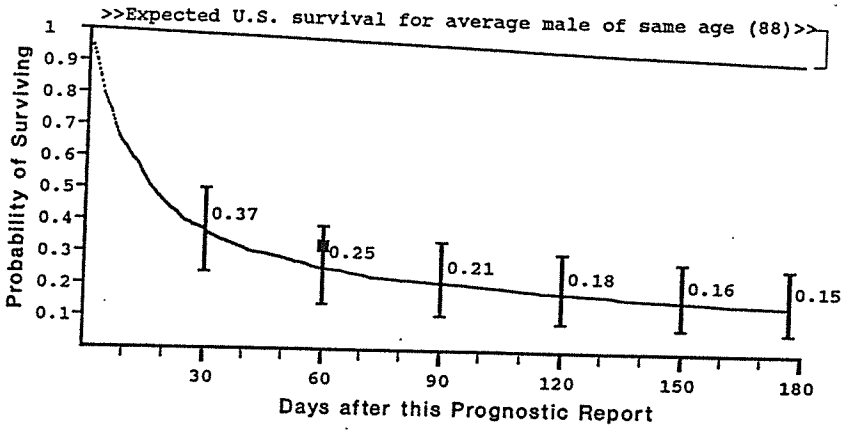
Appendix Figure 4. Effect of previous hospital days on prognosis. Risk for death was associated with the number of days the patient spent in the hospital before study entry. This effect was strongest in patients hospitalized for at least 1 week before study entry.

Prognostic Estimates
Day 3 of Study

Disease GRP: MOSF with Malignancy.

These estimates reflect information available on 07/02/93. Future estimates will change significantly if the patient's physiologic state changes significantly. Disregard previous prognostic estimates.

Estimate of the Probability of Surviving (with confidence bars)



■ Model Estimate enhanced by Physician's Estimate is 0.32
(95% confidence limits are .20 and .45)

Incorporating the physician's estimate into the model can enhance the accuracy of the estimates, since the model doesn't take into consideration all the information that the physician knows about the patient.

RANKING OF VARIABLES IN PROGNOSTIC ESTIMATE
Day 3 of Study

Date that chart data is generated: 07/02/93

Variable	Value	Weight
PaO2/FIO2	99.00	63
HEART RATE	146.00	49
CREATININE	2.10	33
DAYS IN HOSPITAL	13.00	31
CANCER(0/1/2)	1.00	24
TEMPERATURE	36.10	21
WBC	14.70	18
AGE	87.81	14
RESPIRATION RATE	28.00	11
SERUM SODIUM	141.00	0
GLASGOW COMA SCORE	15.00	0
MEANBP	127.00	0
ALBUMIN (Day 3)	2.80	0
BILIRUBIN	0.20 (From an earlier day)	-6

Variables are listed in order by the relative weight they had in estimating this patient's prognosis.

*** Indicates that this value was missing from the chart at the time we attempted to collect this data. A normal value was substituted.

CANCER: 0 = None
1 = Present
2 = Metastatic

Days in Hospital: Number of days in hospital prior to and including study admission date.

Appendix Figure 5. Top. Phase II Feedback report generated on study day 3 for patient admitted to SUPPORT with multiple organ system failure with malignancy. Survival curve is generated for 180 days. This report includes enhancement of model's estimate with the physicians' estimate. **Bottom. Prognostic variables used in generating day 3 feedback report and their relative weights in the estimate for survival.**

tices in acute ischemic heart disease. A prospective multicenter clinical trial. N Engl J Med. 1984;310:1273-8.
21. Goldman L, Cook EF, Brand DA, Lee TH, Rouan GW, Weisberg MC, et al. A computer protocol to predict myocardial infarction in emergency department patients with chest pain. N Engl J Med. 1988;318: 797-803.
22. O'Connor GT, Plume SK, Olmstead EM, Coffin LH, Morton JR, Maloney CT, et al. A regional prospective study of in-hospital mortality associated with coronary artery bypass grafting. The Northern New England Cardiovascular Disease Study Group. JAMA. 1991;266: 803-9.
23. Hannan EL, Kilburn H Jr, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. JAMA. 1990;264:2768-74.
24. Knaus WA, Wagner DP, Lynn J. Short-term mortality predictions for critically ill hospitalized adults: science and ethics. Science. 1991;254: 389-94.
25. Wagner DP, Knaus WA, Harrell FE, Zimmerman JE, Watts C. Daily prognostic estimates for critically ill adults in intensive care units: results from a prospective, multicenter, inception cohort analysis. Crit Care Med. 1994;22:1359-1372.
26. Moskowitz AJ, Kuipers BJ, Kassirer JP. Dealing with uncertainty, risks, and tradeoffs in clinical decisions. A cognitive science approach. Ann Intern Med. 1988;108:435-49.
27. United States. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. Making Health Care Decisions: A Report on the Ethical And Legal Implications Of Informed Consent In The Patient-Practitioner Relationship. Washington D.C.: United States Government Printing Office;1982: 245-6.
28. Cassel CK, Meier DE. Morals and moralism in the debate over euthanasia and assisted suicide. N Engl J Med. 1990;323:750-2.
29. Katz J. The Silent World of Doctor and Patient. New York: Free Press; 1984.
30. Annas GJ. Informed consent, cancer, and truth in prognosis. N Engl J Med. 1994;330:223-5.
31. American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. Crit Care Med. 1992;20:864-74.
32. American Thoracic Society Bioethics Task Force.Withholding and withdrawing life-sustaining therapy. Ann Intern Med. 1991;115:478-85.
33. Schneiderman LJ, Jecker NS, Jonsen AR. Medical futility: its meaning and ethical implications. Ann Intern Med. 1990;112:948-54.
34. Knaus WA, Wagner DP, Zimmerman JE, Draper EA.Variations in mortality and length of stay in intensive care units. Ann Intern Med. 1993;118:753-61.
35. Fleck A, Raines G, Hawker F, Trotter J, Wallace PI, Ledingham IM, et al. Increased vascular permeability: a major cause of hypoalbuminaemia in disease and injury. Lancet. 1985;1:781-4.
36. Austrian R, Gold J. Pneumococcal bacteremia with special reference to bacteremic pneumococcal pneumonia. Ann Intern Med. 1964;60: 759-70.
37. Knaus WA, Harrell FE, Fisher CJ Jr, Wagner DP, Opal SM, Sadoff JC, et al. The clinical evaluation of new drugs for sepsis. A prospective study design based on survival analysis. JAMA. 1993;270:1233-41.
38. Califf RM, Bounous P, Harrell FE, McCants B, Lee KL, McKinnis RA, et al. The prognosis in the presence of coronary artery disease. In: Braunwald E, Mock MB, Watson JT, eds. Congestive Heart Failure: Current Research and Clinical Applications. New York: Grove & Stratton; 1982:31-40.
39. Escarce JJ, Kelley MA. Admission source to the medical intensive care unit predicts hospital death independent of APACHE II score. JAMA. 1990;264:2389-94.
40. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. N Engl J Med. 1985;313: 793-9.
41. Feinstein AR. An additional basic science for clinical medicine: I. The constraining fundamental paradigms. Ann Intern Med. 1983;99:393-7.
42. Knaus WA, Sun X, Nystrom O, Wagner DP. Evaluations of definitions for sepsis. Chest. 1992;101:1656-62.
43. Cowen JS, Kelley MA. Errors and bias in using predictive scoring systems. Crit Care Med. 1994;10:53-72.
44. Jen J, Kim H, Piantadosi S, Liu ZF, Levitt RC, Sistonen P, et al. Allelic loss of chromosome 18g and prognosis in colorectal cancer. N Engl J Med. 1994;331:213-21.
45. Kruse JA, Thill-Baharozian MC, Carlson RW. Comparison of clinical assessment with APACHE II for predicting mortality risk in patients admitted to a medical intensive care unit. JAMA. 1988;260:1739-42.
46. McClish DK, Powell SH. How well can physicians estimate mortality in a medical intensive care unit? Med Decis Making. 1989;9:125-32.
47. Brannen AL, Godfrey LJ, Goetter WE. Prediction of outcome from critical illness. A comparison of clinical judgement with a prediction rule. Arch Intern Med. 1989;149:1083-6.