Intensive Care Med (2004) 30:655–659
DOI 10.1007/s00134-003-2139-7

ORIGINAL

Fábio Gusmão Vicente
Frederico Polito Lomar
Christian Mélot
Jean-Louis Vincent

Can the experienced ICU physician predict ICU length of stay and outcome better than less experienced colleagues?

Received: 25 February 2003
Accepted: 12 December 2003
Published online: 21 January 2004

F. Gusmão Vicente · F. Polito Lomar · C. Mélot · J.-L. Vincent (✉)
Department of Intensive Care Unit,
Erasme Hospital Free University of Brussels,
Route de Lennik 808, 1070 Brussels, Belgium
e-mail: jlvincen@ulb.ac.be
Tel.: +32-2-5553380
Fax: +32-2-5554555

Abstract *Objective:* To assess the ability of physicians with varying degrees of experience to predict the length of stay and outcome of intensive care unit (ICU) patients. *Design:* Prospective, interview-based study. *Setting:* A 31-bed mixed medical-surgical ICU. *Patients:* A total of 223 consecutive patients (excluding those admitted for routine post-operative surveillance) admitted to the ICU. *Interventions:* None. *Measurements and main results:* Physicians immediately responsible for each patient, and others fully aware of the case, were interviewed separately during the first 12 h of ICU admission to determine their assessment of the patient's likely duration of stay on the ICU and the probable outcome. Degree of predictive accuracy was assessed using the Kappa statistic with kappa $\leq$0.2 poor, 0.21–0.4 fair, 0.41–0.60 moderate, 0.61–0.8 good, and 0.81–1.0 very good. Physicians were graded according to their degree of experience as junior (less than 1 year ICU experience), medium (critical care fellow), and senior (staff physician with supervising functions). For lengths of stay less than 5 days, senior physicians were better predictors than less experienced doctors. For outcome prediction, physicians were generally moderately good at predicting death, with senior physicians tending to be more accurate than their less experienced colleagues (senior kappa 0.68, medium kappa 0.52, junior kappa 0.43). *Conclusions:* Prediction of length of ICU stay was poor amongst all physicians in patients with a length of stay greater than 5 days. Experienced physicians were better predictors of ICU lengths of stay less than 5 days and, in contrast to some reports, of ICU outcome than their more inexperienced counterparts.

Keywords Predictive value · Organ failure · Mortality · Survival

Introduction

Outcome prediction is becoming an increasingly important area of clinical intensive care medicine as costs continue to rise and patients, families, physicians, and managers demand more informed health care decisions. Such decisions need to be employed early in, or even before, the admission process to identify which patients are likely to benefit from expensive intensive care unit (ICU) facilities and treatment, and who almost certainly will not. Once on the ICU, given the often acute nature of critical care, the intensivist frequently needs to make rapid decisions regarding treatment and procedures; such judgments will again often be based on an estimate of how the patient is likely to respond and what the prognosis is likely to be. In addition, perhaps the most common request from a patient and/or their relatives is when they will be able to go home or, indeed, whether they will "pull through" at all.

The likely length of stay of a patient may also influence therapeutic decisions. Several recent studies have indicated that some therapeutic strategies that

impact on patient outcome may only have an effect on patients with longer ICU stays. For example, intensive glycemic control reduced morbidity and mortality in a surgical ICU, but primarily in patients who stayed in the ICU for more than 5 days [1], and recombinant erythropoietin administration has been shown to increase hematocrit and reduce blood transfusions, but at least 5 days are needed to produce these effects [2]. Offering these therapies to patients who are unlikely to stay on the ICU for more than a few days may, therefore, be an unnecessary use of resources.

Several objective systems have been developed to predict mortality in the ICU patient, including the acute physiology, age, and chronic health evaluation (APACHE) score [3], and the simplified acute physiology score (SAPS) [4]. Most physicians will make an almost instant, subjective, and often subconscious, assessment of a patient's likely outcome during their first meeting, and some studies have suggested that such medical judgments may be equal, or even superior, to prognostic indexes (whose observed variables do not take into account the subjective impression of the physician) usually employed in the ICU environment [5, 6, 7, 8, 9, 10, 11, 12, 13]. Prediction of length of stay is perhaps more difficult and has been less studied. Tu and Mazer noted that experienced physicians were able to accurately predict which patients would have short lengths of stay after cardiac surgery but longer stay patients were often not identified correctly [14].

The present study aimed to evaluate the ability of ICU physicians to predict the length of stay and outcome of critically ill patients, and in addition, hypothesized that the accuracy of prediction would be influenced by the degree of medical experience of the predictor.

Material and methods

The study was conducted in a tertiary care teaching hospital on a 31-bed medical-surgical ICU. Data were collected prospectively from 1st May to 30th June 2002. The inclusion criteria were: all patients over 15 years old admitted to the unit for medical reasons, urgent surgical treatment, or expected post-operative complications after scheduled surgery. Exclusion criteria were: patients whose total stay was less than 12 h and patients admitted for routine surveillance after scheduled surgery whose length of stay was expected to be no more than 24 h. Transplanted patients were included as non-scheduled surgery.

During the first 12 h after admission, a brief interview was conducted with the physicians, asking about the expected length of stay (in days) and the ICU outcome (death versus discharge). The interview was performed with the doctors immediately responsible for the patient's care and with other members of the team when fully aware of the case. We graded the physicians as follows: junior—less than 1 year of experience in critical care, medium—critical care fellow and senior—staff physician with supervising functions. The interviews were conducted separately to ensure that the members of the team were not aware of each other's opinion and with a short space of time (less than 30 min) between interviews for the same patient in order to ensure the information known by one physician was the same as that known by the others.

Demographic data (age, gender), date of arrival on the ICU, number of co-morbidities before admission, origin of the patient (emergency room or ward), the main reason why the patient was admitted to the ICU (medical or surgical problem), primary ICU diagnosis, 24-h SAPS II score, need of invasive mechanical ventilation or renal replacement therapy, and use of vasopressor drugs (dopamine and/or norepinephrine) were recorded.

The research physicians visited the admitted patients on a daily basis throughout their ICU stay, including during the weekends, until death or discharge. The length of stay was counted in days beginning at midnight [15]. When discharge was postponed for non-medical organizational reasons, the day on which discharge was decided was the day taken into account.

The data used to complete the protocol were available on a computer-based system of the ICU, but the team was not necessarily aware of the sum of the SAPS-II or its predicted mortality index.

Statistical analysis was done using the Kappa statistic for the agreement between the real length of stay and the physician's estimation, dichotomizing a continuous variable. With regard to the outcome prediction, we also employed the Kappa statistic. The values of the Kappa statistic and the strength of agreement were considered as follows: $\leq$0.2 poor, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good and 0.81–1.0 very good [16].

For comparisons related to other variables we used Student's *t*-test and the Fisher exact test with a confidence interval of 95% and a $p<0.05$ considered as significant. The statistical packages used were Analyse-it version 1.63 for Microsoft Excel (Analyse-It Software, UK), SPSS version 10.0 for windows (SPSS, Chicago, Ill) and Medcalc version 6.16.

Results

We enrolled 223 patients over the 2-month period and performed a total of 318 interviews (127 from 6 senior doctors, 107 from 5 medium grade doctors and 84 from 7 junior doctors). Table 1 summarizes the clinical data of the patients included in the study. The population studied had a mean age of 59 years. The mean SAPS-II index was 33. The mean length of stay was 6.8 days with a median of 3 days. The mortality rate was 18.0%.

The SAPS II score was 51 among the patients who died in the ICU versus 29 among the patients discharged to wards ($p<0.001$). There was a trend towards a longer length of ICU stay in patients who came from the ward than from the emergency room.

Length of stay

The Kappa statistic for the prediction of length of stay in the ICU, according to the degree of medical experience, versus the true length of stay is shown in Fig. 1. Senior physicians had good predictive accuracy for lengths of stay less than or equal to 4 days (Kappa $\geq$0.61), moderate predictive accuracy for lengths of stay of 5 and 6 days (Kappa between 0.60 and 0.41), and fair predictive accuracy for lengths of stay greater than 7 days (Kappa <0.40). For medium grade physicians, predictive accuracy

Table 1 Clinical data of the patients

Total patients		Number	%
		223	100
Gender	Male	140	62.8
	Female	83	37.2
Mean age (years)		59	(range 15–93)
Mean SAPS II		33	(range 5–91)
Mean (median) length of stay (days)		6.8 (3.0)	(range 1–70)
Outcome	Death	40	18
	Discharge	183	82
Patient origin	ER	135	60.5
	Ward	88	39.5
Problem	Medical	162	72.6
	Surgical	61	27.4
Co-morbidities	Arterial hypertension	69	30.9
	Diabetes mellitus	38	17
	Cerebrovascular disease	16	7.2
	Heart disease	79	35.4
	Respiratory disease	41	18.4
	Liver disease	16	7.2
	Renal disease	24	10.8
	History of malignancy	16	7.2
	Other	68	30.8
Primary ICU diagnosis	Respiratory failure	47	21.1
	Shock	35	15.7
	Coma	35	15.7
	Trauma	26	11.7
	Infection/sepsis	23	10.3
	Cardiorespiratory arrest	5	2.2
	Other	52	23.3

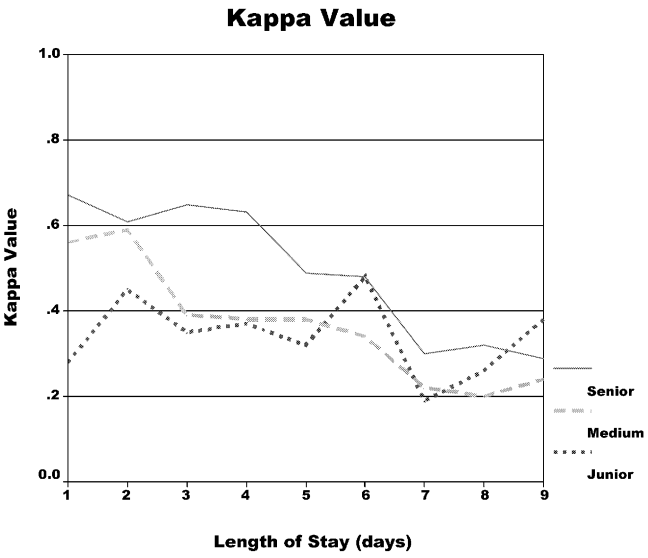

Fig. 1 Kappa value for length of stay according to degree of physician experience

was moderate for lengths of stay of 1 or 2 days and fair for longer lengths of stay. For junior physicians there was no trend in predictive ability with length of stay as for more senior colleagues; predictive accuracy was fair or poor (Kappa <0.40) for all lengths of stay except for lengths of stay of 2 and 6 days when the accuracy was moderate (Kappa between 0.60 and 0.41).

The use of mechanical ventilation or the presence of shock on admission to the ICU was associated with a significantly longer mean length of stay in the ICU (Table 2). In addition, patients who needed mechanical

Table 2 Comparisons between mean length of stay and outcome with other variables

Variables	No of patients	Mean LOS	*p*-value	Survival	*p*-value
Gender	Male (140)	6.95	0.77	82.8	0.82
	Female (83)	6.58		80.7	
Origin	ER (135)	5.86	0.053	80.7	0.65
	Ward (88)	8.27		84.1	
MV on admission	Yes (86)	10.31	<0.001	66.3	<0.001
	No (137)	4.61		95	
RRT on admission	Yes (9)	9.55	0.357	55.5	0.11
	No (214)	6.70		83.2	
Shock on admission	Yes (57)	10.28	<0.001	72	0.04
	No (166)	5.62		85.5	
MV during stay	Yes (18)	23.33	<0.001	55.5	0.01
	No (205)	5.36		84.4	
RRT during stay	Yes (11)	15.27	<0.002	90.9	0.76
	No (212)	6.37		81.6	
Shock during stay	Yes (17)	15.88	<0.001	64.7	0.12
	No (206)	6.06		83.5	
SAPS II	≥25 (152)	7.58	0.052	75.7	<0.001
	<25 (71)	5.17		95.8	
Outcome	Death	6.725	0.947		
	Discharge	6.831			

ER Emergency room.
RRT Renal replacement therapy.
MV Mechanical ventilation.

Table 3 Kappa value for outcome prediction

Grade	n	Kappa	95% CI
Senior	127	0.68	0.49–0.87
Medium	107	0.52	0.31–0.72
Junior	84	0.43	0.15–0.71
Overall	318	0.55	0.43–0.68

ventilation, renal replacement therapy, or vasopressor agents during their ICU stay had a significantly longer mean length of stay than the other patients.

Outcome

Overall, physicians underestimated mortality. In the 318 interviews conducted, 48 of those interviewed (15%) felt that the patient would die and 270 (85%) that the patient would be discharged, while the real outcome was that 63 (20%) patients died and 255 (80%) were discharged (considering one answer for one patient). The overall Kappa value for outcome prediction was 0.55, with the senior physicians showing a trend towards a better predictive accuracy than their less experienced colleagues (Table 3).

Patients receiving mechanical ventilation on admission or during their stay, and those with circulatory failure on admission had higher mortality rates (Table 2).

Discussion

The ability to predict length of stay and likely outcome may be important for therapeutic and logistic reasons. Predictions of outcome, both in terms of length of stay and mortality, can assist in the planning and optimization of intensive care facilities and utilization. ICU admission should be restricted to those likely to benefit most from the care provided [17]; clearly a patient unlikely to survive for more than a few hours or days may not be the most appropriate candidate for an ICU bed. Similarly, predictions of outcome may be used to decide when to discharge, when to withdraw or withhold therapy, and even when not to resuscitate [18].

Therapeutic decision-making may also be influenced by predictions regarding likely length of stay and outcome. For example, in the recent study by Van den Berghe and colleagues [1], intensive glycemic control in a surgical ICU population reduced morbidity and mortality primarily among patients who remained in the ICU for more than 5 days. Similarly, the administration of recombinant human erythropoietin to critically ill patients is effective in raising their hematocrit concentration and in reducing the total number of units of red blood cells they require, but at least 5 days treatment is needed to produce this effect [2]. Targeting such treatments at patients who are unlikely to survive beyond a few days may thus be an unnecessary, and even wasteful, use of already limited resources.

Our results showed that the ability of physicians to predict the length of stay of a patient in the ICU is not precise beyond 4 days, and with increasing ICU length of stay beyond this, prediction accuracy worsened. For an ICU stay less than 5 days, prediction ability was associated with degree of ICU experience, with senior physicians being better able to estimate length of stay than their less experienced counterparts. For junior physicians, there was no decline in Kappa values in parallel with increasing length of stay as observed for the more senior physicians, suggesting that their prediction was more a random guess than an informed opinion.

In a Canadian study of patients undergoing cardiac surgery [14], experienced clinicians preoperatively predicted which patients would have a short ICU length of stay ($\leq$2 days) with a considerable degree of accuracy (correct prediction in 87.6% of patients). However, as in our study, for longer lengths of stay predictive ability was poor (only 39.4% of the patients with an ICU stay of greater than 2 days were identified) [14].

Prediction of outcome can alter patient management, with increased use of certain aspects of intensive care in patients expected to have a good outcome and reduced use of these same resources in patients not expected to survive [19]. In our study, although not statistically significant, there was a trend towards improved estimating ability with greater experience with senior physicians having a good capacity to predict outcome, and medium grade and junior physicians having a moderate ability. Others have also reported improved predictive accuracy with greater degrees of experience. Christensen et al. [5] noted better discrimination of mortality prediction by the more experienced attending physician than by less qualified fellows or residents, but only within the first 24 h of ICU stay. Barrera and colleagues [20] reported that intensive care physicians made more accurate predictions than primary team physicians when assessing the outcome of critically ill patients, and Copeland-Fields et al. [13] found that intensive care physicians were more accurate than emergency physicians who were in turn better than intensive care nurses in predicting survival, but all three groups were better than the mortality prediction model (MPM). However, Kruse et al. [6] found no difference in the ability of various grades of physician (fellows, residents, and interns) to predict survival, although physicians were more accurate than nurses.

Objective grading scales are not designed to be applied to individual patients; although highly specific (i.e., good at predicting survival) they are not so sensitive (i.e., not good at predicting death) and cannot be relied on to predict outcome for individual patients [21]. Indeed,

many studies have shown subjective estimates of mortality risk to be as, or more, accurate than those based on such scoring systems. In a group of 568 patients admitted to a general ICU and with a mortality rate of 45.8% (260 deaths), Marks et al. [8] found that the subjective prediction of doctors and nurses was a more powerful predictor of outcome than the APACHE II predicted risk of death, and Kruse and colleagues [6] showed that the estimation of mortality risk by physicians or nurses was at least as good as the APACHE II predictions in 366 medical ICU patients. Although we did not compare a predictor score with clinical judgment, our study showed that experienced physicians were able to predict outcome with a moderate degree of accuracy.

Elective surgical patients were excluded from the study to limit the potential over-estimation of the results with the inclusion of these patients as most of them were expected to stay for only 24 h, i.e., their length of stay would have been easily and precisely predicted by the doctors. In addition, the physician's predictive accuracy for the length of stay for a similar group of patients (post-operative cardiac surgery patients) has been well described [14].

We acknowledge some possible drawbacks to this type of study. Firstly, the numbers are relatively small which necessarily limits the interpretations that can be drawn. Secondly, it was sometimes difficult to gather evaluations from several doctors, especially during the weekends. Also, when senior physicians were attending a patient, medium and junior physicians were not always fully aware of the case, therefore their predictions could not be considered. Finally, sometimes we could not locate the attending physician in the short period of time (30 min) available. Nevertheless, despite these potential limitations, the study provides interesting data with implications for treatment and research.

In conclusion, outcome prediction is not very accurate, but experienced intensive care physicians are better predictors of ICU lengths of stay less than 5 days and of ICU outcome than their more inexperienced colleagues.

References

1. Van den Berghe G, Wouters P, Weekers F et al. (2001) Intensive insulin therapy in the critically ill patient. N Engl J Med 345:1359–1367
2. Corwin HL, Gettinger A, Rodriguez RM et al. (1999) Efficacy of recombinant human erythropoietin in the critically ill patient: a randomized, double-blind, placebo-controlled trial. Crit Care Med 27:2346–2350
3. Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985) APACHE II: a severity of disease classification system. Crit Care Med 13:818–829
4. Le Gall J-R, Lemeshow S, Saulnier F (1993) A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. JAMA 270:2957–2963
5. Christensen C, Cottrell JJ, Murakami J, Mackesy ME, Fetzer AS, Elstein AS (1993) Forecasting survival in the medical intensive care unit: a comparison of clinical prognoses with formal estimates. Methods Inf Med 32:302–308
6. Kruse JA, Thill-Baharozian MC, Carlson RW (1988) Comparison of clinical assessment with APACHE II for predicting mortality risk in patients admitted to a medical intensive care unit. JAMA 260:1739–1742
7. Brannen AL, Godfrey LJ, Goetter WE (1989) Prediction of outcome from critical illness. A comparison of clinical judgment with a prediction rule. Arch Intern Med 149:1083–1086
8. Marks RJ, Simons RS, Blizzard RA, Browne DR (1991) Predicting outcome in intensive therapy units—a comparison of Apache II with subjective assessments. Intensive Care Med 17:159–163
9. Knaus WA, Harrell FE Jr, Lynn J et al. (1995) The SUPPORT prognostic model. Objective estimates of survival for seriously ill hospitalized adults. Study to understand prognoses and preferences for outcomes and risks of treatments. Ann Intern Med 122:191–203
10. McClish DK, Powell SH (1989) How well can physicians estimate mortality in a medical intensive care unit? Med Decis Making 9:125–132
11. Moreno R, Apolone G, Miranda DR (1998) Evaluation of the uniformity of fit of general outcome prediction models. Intensive Care Med 24:40–47
12. Woods AW, MacKirdy FN, Livingston BM, Norrie J, Howie JC (2000) Evaluation of predicted and actual length of stay in 22 Scottish intensive care units using the APACHE III system. Acute physiology and chronic health evaluation. Anaesthesia 55:1058–1065
13. Copeland-Fields L, Griffin T, Jenkins T, Buckley M, Wise LC (2001) Comparison of outcome predictions made by physicians, by nurses, and by using the mortality prediction model. Am J Crit Care 10:313–319
14. Tu JV, Mazer CD (1996) Can clinicians predict ICU length of stay following cardiac surgery? Can J Anaesth 43:789–794
15. Marik PE, Hedman L (2000) What's in a day? Determining intensive care unit length of stay. Crit Care Med 28:2090–2093
16. Mould RF (1998) Introductory medical statistics. Institute of Physics Publishing, Bristol
17. Society of Critical Care Medicine Ethics Committee (1994) Consensus statement on the triage of critically ill patients. JAMA 271:1200–1203
18. Silverstein MD (1988) Prediction instruments and clinical judgment in critical care. JAMA 260:1758–1759
19. Murray LS, Teasdale GM, Murray GD et al. (1993) Does prediction of outcome alter patient management? Lancet 341:1487–1491
20. Barrera R, Nygard S, Sogoloff H, Groeger J, Wilson R (2001) Accuracy of predictions of survival at admission to the intensive care unit. J Crit Care 16:32–35
21. Suter PM, Armaganidis A, Beaufils F et al. (1994) Consensus conference organized by the ESICM and SRLF: predicting outcome in ICU patients. Intensive Care Med 20:390–397