Journal of Critical Care (2008) **23**, 339–348

ELSEVIER

Journal of Critical Care

Modeling in-hospital patient survival during the first 28 days after intensive care unit admission A prognostic model for clinical trials in general critically ill patients☆

Rui P. Moreno MD, PhD[a,*], Philipp G.H. Metnitz MD, PhD, DEAA[b], Barbara Metnitz MS, PhD[c], Peter Bauer PhD[c], Susana Afonso de Carvalho MD[a], Anette Hoechtl PhD[b] on behalf of the SAPS 3 Investigators

[a]*Unidade de Cuidados Intensivos Polivalente, Hospital de St. António dos Capuchos, Centro Hospitalar de Lisboa Central, E. P. E., Lisboa, Portugal*
[b]*Department of Anaesthesiology and General Intensive Care, University Hospital of Vienna, Vienna, Austria*
[c]*Medical University of Vienna, Core Unit for Medical Statistics and Informatics, Section of Medical Statistics, Vienna, Austria*

Keywords:
Intensive care;
Critical care;
Severity scores;
Outcome;
28-day survival

Abstract

Objective: The objective of the study was to develop a model for estimating patient 28-day in-hospital mortality using 2 different statistical approaches.

Design: The study was designed to develop an outcome prediction model for 28-day in-hospital mortality using (*a*) logistic regression with random effects and (*b*) a multilevel Cox proportional hazards model.

Setting: The study involved 305 intensive care units (ICUs) from the basic Simplified Acute Physiology Score (SAPS) 3 cohort.

Patients and Participants: Patients (n = 17 138) were from the SAPS 3 database with follow-up data pertaining to the first 28 days in hospital after ICU admission.

Interventions: None.

Measurements and Results: The database was divided randomly into 5 roughly equal-sized parts (at the ICU level). It was thus possible to run the model-building procedure 5 times, each time taking four fifths of the sample as a development set and the remaining fifth as the validation set. At 28 days after ICU admission, 19.98% of the patients were still in the hospital. Because of the different sampling space and outcome variables, both models presented a better fit in this sample than did the SAPS 3 admission score calibrated to vital status at hospital discharge, both on the general population and in major subgroups.

☆ Authors' contribution and competing interests. Rui Moreno and Philipp Metnitz actively organized and chaired the SAPS 3 project (see electronic supplementary material for a complete list of participants in the project) and actively participated in all steps of data collection, analysis, and model development. Barbara Metnitz and Peter Bauer were responsible for data management and statistical analysis of the SAPS 3 project. Manuscript preparation was done by Rui Moreno, Philipp Metnitz, and Barbara Metnitz.

* Corresponding author. Fax: +351 21 3153784.
E-mail address: r.moreno@mail.telepac.pt (R.P. Moreno).

0883-9441/$ – see front matter
doi:10.1016/j.jcrc.2007.11.004

Conclusions: Both statistical methods can be used to model the 28-day in-hospital mortality better than the SAPS 3 admission model. However, because the logistic regression approach is specifically designed to forecast 28-day mortality, and given the high uncertainty associated with the assumption of the proportionality of risks in the Cox model, the logistic regression approach proved to be superior.

1. Introduction

All general outcome prediction models in intensive care aim at predicting vital status at hospital discharge on the basis of a given set of variables collected at admission or shortly thereafter [1,2]. However, regulatory agencies such as the US Food and Drug Administration often rely mainly on outcome data pertaining to vital status at 28 days after intensive care unit (ICU) admission. This holds especially true for clinical studies in which the statistical model used should be able to control for differences in severity among different groups of patients with different outcomes or assigned to different treatments. Survival at 28 days after ICU admission is thus the most frequently used end point in randomized clinical trials conducted in critically ill patients, for example, those studying new therapeutic approaches for patients with sepsis or septic shock. For that reason, a prognostic tool that is able to model survival during this limited time frame would be very useful in comparing the ratio of observed to expected mortality in different groups of patients, stratified by one intervention. Until now, no specific instrument has been developed and published for this purpose.

Moreover, current outcome prediction models were built almost always by logistic regression. Such models use a series of variables, collected at the patient level, that measure the degree of physiological reserve (age, comorbid diseases), the reasons for and circumstances of ICU admission, and the presence and degree of organ dysfunction to model vital status at hospital discharge [1,2].

Crude mortality rates are the most commonly used estimators of mortality; the Kaplan-Meier estimator is used usually only to display the distribution of survival times. Logistic regression models are widely used to define factors related to the probability of death, whereas Cox models are more often used to model the instantaneous risk of nonfatal outcomes [3]. Whether one of these 2 models is more appropriate in identifying prognostic factors in intensive care is the subject of intense debate [4].

The objective of this study was thus to develop a model to predict 28-day in-hospital mortality using 2 different statistical approaches.

2. Materials and methods

2.1. Project organization and data collection

The SAPS 3 project was conducted by the SAPS 3 Outcomes Research Group. The project was endorsed by the European Society for Intensive Care Medicine (ESICM) (www.esicm.org) and conducted in cooperation with the Section on Health Services Research and Outcome of the ESICM. The organization of the study, methods used for data collection, and details about the variables collected have been described elsewhere [5].

Data were collected at ICU admission; on days 1, 2, and 3; and on the last day of the ICU stay. All patients were subject to mandatory follow-up until 28 days after ICU admission or until hospital discharge, whichever came first. Patients remaining in the hospital at 90 days were classified as being "still in the hospital." Data were collected from all consecutively admitted patients between 14 October and 15 December 2002.

2.2. Statistical analysis

Statistical analysis was performed using the SAS system, version 9.1 (SAS Institute Inc, Cary, NC). A P value < .05 was considered significant. Unless otherwise specified, results are expressed as median and quartiles. Observed-to-expected mortality ratios were calculated by dividing the number of observed deaths per group by the number of expected deaths per group (as predicted by the SAPS II). To test for statistical significance, we calculated 95% confidence intervals according to the method described by Hosmer and Lemeshow [6].

For development of the SAPS 3 28-day score, 2 strategies were applied (1) to model vital status at 28 days as a binary outcome (alive/dead), fitting an outcome prediction model using logistic regression with random effects, and (2) to develop a multilevel proportional hazards model, where all patients discharged from hospital were kept in the risk set for 28-day in-hospital mortality (modeling transition probabilities according to the subdistribution of Fine and Gray [7]).

In both cases, for cross-validation of the models, the database was divided randomly into 5 roughly equal-sized parts (at ICU level). It was thus possible to run the model-building procedure 5 times, each time taking four fifths of the sample as a development set and the remaining fifth as the validation set. The Hosmer-Lemeshow goodness-of-fit $\hat{H}$ statistic and $\hat{C}$ statistic [6] were used to evaluate the calibration of the model. Discrimination was tested by measuring the area under the receiver operating characteristic (aROC) curve, as described by Hanley and McNeil [8]. Expected survival curves in subgroups were calculated by averaging predicted survival curves for the individual patients in the subgroups.

In both cases, multilevel modeling (also called *hierarchical modeling*) has been applied to account for the variations among centers.

3. Results

From the basic SAPS 3 cohort (19 577 patients), we included 17 138 patients (from 305 ICUs) for whom there were follow-up data pertaining to the first 28-days in-hospital after ICU admission (admission-, discharge-, and outcome-data for these patients are presented in Tables 1 and 2). At 28 days after ICU admission, 19.98% of the patients were still in the hospital.

Not surprisingly, the SAPS 3 admission score (which was developed for a different outcome variable, namely, vital status at hospital discharge) could not reliably predict vital status at day 28 after ICU admission, either for the global cohort (Hosmer-Lemeshow goodness-of-fit test $\hat{H}$, 179.86 [$P < .01$]; Hosmer-Lemeshow goodness-of-fit test $\hat{C}$, 173.29 [$P < .01$]; aROC curve, 0.837) or for major subgroups (electronic supplementary material [ESM], Appendix L).

Table 1 The ICU admission data for the SAPS 3 28-day cohort (n = 17 138)

Patient characteristics	n	%
Sex		
Female	6735	39.3
Male	10 390	60.6
Missing data	13	0.1
Age (y) (median, Q1-Q3)	64 (49-74)	
Initial location		
Home	2398	14.0
Same hospital	12 294	71.7
Chronic care facility	66	0.4
Public place	441	2.6
Other hospital	1846	10.8
Other	60	0.4
Missing data	33	0.2
Intrahospital location before ICU admission		
Emergency room	4755	27.7
Intermediate care unit/High-dependency unit	494	2.9
Operative room	6559	38.3
Other	423	2.5
Other ICU	625	3.6
Recovery room	397	2.3
Ward	3086	18.0
Missing data	799	4.7
ICU admission status		
Planned	5688	33.2
Unplanned	11 060	64.5
Missing	390	2.3
Acute infection at ICU admission		
No infection	13 223	77.2
Clinically improbable/colonization	300	1.8
Clinically probable/documented	2488	14.5
Microbiologically documented	1114	6.5
Missing data	13	0.1
Surgical status		
No surgical procedure	7474	43.6
Scheduled surgery	5796	33.8
Emergency surgery	3011	17.6
Missing data	857	5.0

Q1-Q3 indicates first to third quartiles.

Table 2 The ICU discharge and outcome data for the SAPS 3 28-day cohort (n = 17 138)

Patient characteristics	n	%
ICU LOS (d) (median, Q1-Q3)	2 (1-6)	
ICU discharge destination		
Home	372	2.2
Same hospital	12 776	74.5
Other hospital	887	5.2
Missing data	3103	18.1
Intrahospital discharge		
Emergency room	51	0.3
Intermediate care unit/ High-dependency unit	1966	11.5
Other	262	1.5
Other ICU	504	2.9
Recovery room	220	1.3
Ward	10 483	61.2
Missing data	3652	21.3
ICU discharge status		
Planned discharge	12 590	73.5
Unplanned discharge	1480	8.6
Missing data	3068	17.9
Risk adjustment		
SOFA score (median, Q1-Q3)	9 (7-11)	
Outcome		
ICU mortality		17.4

LOS indicates length of stay.

3.1. Modeling 28-days after ICU admission in-hospital vital status as a binary outcome (alive/dead)

To estimate vital status (alive or dead) of patients at 28 days after ICU admission, we developed an outcome prediction model using logistic regression with random effects. All the procedures for data handling were those used in the development of the SAPS 3 admission score as described before [2]. The final score sheet is presented in Tables 3 and 4, and the corresponding estimated and P values are given in the ESM, Appendix D. The relationship between the SAPS 3 28-day score and in-hospital vital status at 28 days after ICU admission is given by the following equation:

$$\text{Logit} = -26.2477 + \ln(\text{SAPS 3 28-day score} + 4.5973) \times 6.0521$$

Table 3 The SAPS 3 28-day score sheet of the multilevel logistic regression model

Part 1											
Box I	4	0	3	4	5	6	7	8	11	14	17
Age (y)		<40			≥40-<60			≥60 <70	≥70 <75	≥75 <80	≥80
Comorbidities			Cancer therapy				Chronic heart failure (NYHA IV)	Cirrhosis, AIDS [a]	Cancer		
BMI	<18.5	≥18.5 <14									
LOS before ICU admission (d)				≥14							
Intrahospital location before ICU admission					Emergency room		Other ICU, other [a]				
Use of major therapeutic options before ICU admission			Vasoactive drugs								
Box II				0	2	3	4	6			
ICU admission: planned or unplanned						Unplanned					
Reason(s) for ICU admission	See part 2										
Surgical status at ICU admission				Scheduled surgery				No surgery, emergency surgery [a]			
Anatomical site of surgery	See part 2										
Acute infection at ICU admission						Pneumonia	Nosocomial				

Box III	14	10	9	7	6	5	3	2	0	2	3	4	6	7
Estimated Glasgow Coma Scale (lowest) (points)	3-4	5		6					≥7					
Total bilirubin (highest) (mg/dL)									<2		≥2			
Body temperature (highest) (°C)					<35				≥35					
Creatinine (highest) (mg/dL)									<1.2	≥1.2-<2				≥2
Heart rate (highest) (beats/min)									<120			≥120-<160	≥160	
Leukocytes (lowest) (g/L)									<15	≥15				
Hydrogen ion concentration (highest) (pH)								<7.30	≥7.30					
Platelets (lowest) (g/L)	<20			≥20-<50		≥50-<100			≥100					
Systolic blood pressure (lowest) (mm Hg)		<40	≥40-<70				≥70-<120		≥120					
Oxygenation		Pao_2/Fio_2 <100 and MV		Pao_2 <60 and no MV Pao_2/Fio_2 ≥100 and MV					Pao_2 ≥60 and no MV					

NYHA indicates New York Heart Association; BMI, body mass index.

Table 4 The SAPS 3 28-day score sheet of the multilevel logistic regression model

Box II	Points
ICU admission [a]	20
Reason(s) for ICU admission	
Cardiovascular	
Rhythm disturbances	−5
Hypovolemic hemorrhagic shock, hypovolemic nonhemorrhagic shock	4
Septic shock	5
Anaphylactic and mixed and undefined shock	6
Cardiogenic shock	8
Neurologic	
Focal neurologic deficit (hemiplegia, paraplegia, tetraplegia)	5
Intracranial mass effect	10
Hepatic	
Liver failure	6
Digestive	
Severe pancreatitis	7
Type and anatomical site of surgery	
Transplantation: liver, kidney, pancreas, kidney and pancreas, other	−15
Trauma: multiple	−12
Trauma: other, isolated (includes thorax, abdomen, limb)	−5
Heart surgery: CABG without valvular repair	−4
Neurosurgery: cerebrovascular accident	5
Upper gastrointestinal surgery	4

CABG, coronary artery bypass graft.
[a] Every patient gets an offset of 20 points for being admitted (to avoid negative scores).

and the probability of in-hospital mortality at the 28 days after ICU admission is given by the following equation:

$$\text{Probability of death} = \frac{e^{\text{logit}}}{1 + e^{\text{logit}}}.$$

The relationship between the SAPS 3 28-day score and the probability of death in the hospital is shown in Fig. 1. Overall, deviations between observed and expected outcomes across all of the strata were not outside sampling variability, as demonstrated by a Hosmer-Lemeshow goodness-of-fit test $\hat{H}$ of 14.03 (P = .17) and a Hosmer-Lemeshow goodness-of-fit test $\hat{C}$ of 9.39 (P = .50). Calibration curves in major patient subgroups are presented in the ESM, Appendix E.

The overall discriminatory capability of the model, as measured by the aROC curve, was 0.837, very similar to the SAPS 3 admission score [2].

Customized equations for each major geographical area are presented in the ESM, Appendix F. Observed-to-expected mortality ratios per geographical area (Appendix G) and per country (Appendix H), together with the respective calibration curves and aROC curves (Appendix I and J), can also be found in the ESM.

3.2. Modeling survival curves by developing a multilevel Cox proportional hazards model using the Fine and Gray technique

Using the proportional hazards model, we fitted a multivariate multilevel model. The coefficients are presented in Appendix K of the ESM. The score sheet for the multiplicative exponential predictor in the hazard function developed by this technique is presented in Appendix L. The expected survivor function cannot be calculated without giving the estimates of the baseline hazard.

However, the degree of correspondence between the observed (in our situation, identical to the Kaplan-Meier estimate) and expected survival curves is shown in Fig. 2. Other curves for major subgroups can be found in the ESM, Appendix M. As for the logistic regression with random effects model, the fit was reasonable in all the development and validation samples and in all patient subgroups.

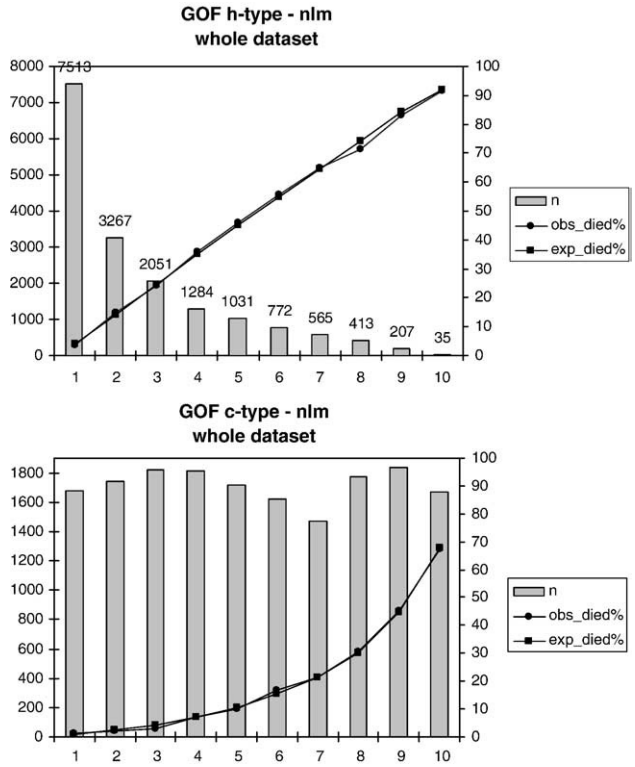

Fig. 1 Goodness-of-fit for the multilevel logistic regression model in the whole data set. Top, Patients are divided into 10 equal groups according to the estimated risk of death ($\hat{H}$ test); test statistic: 5.83, P = .83. Bottom, Patients are divided into 10 equal groups according to the number of patients ($\hat{C}$ test); test statistic: 13.28, P = .21.

The survivor rates obtained by multilevel logistic regression, the Cox proportional hazards model, and Kaplan-Meier estimates in the general population and in major subgroups are compared in Table 5.

4. Discussion

All currently used outcome prediction models in intensive care use a series of variables (collected at the

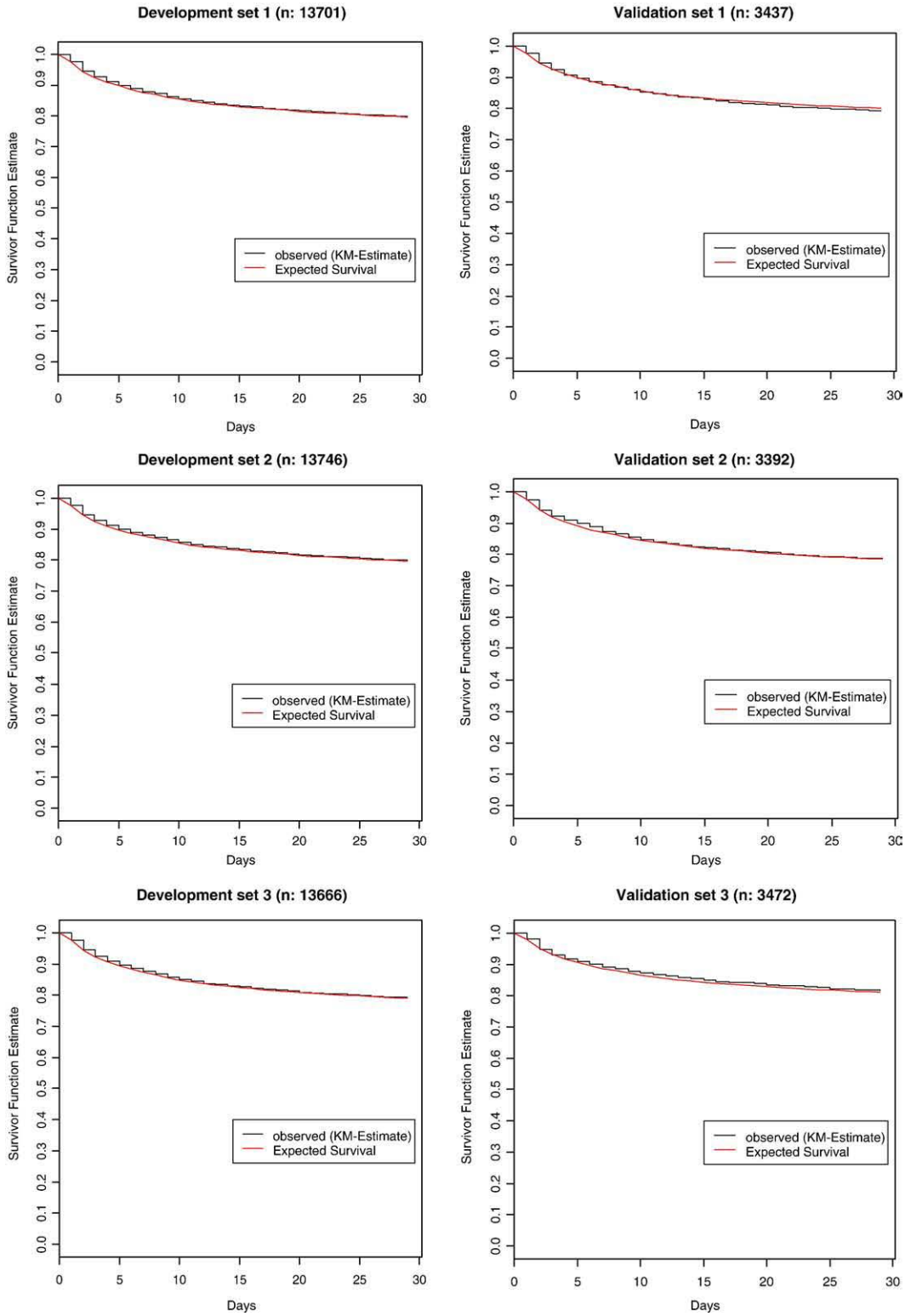

Fig. 2 Degree of correspondence between the observed and expected survival curves in the 5 development and validation samples.

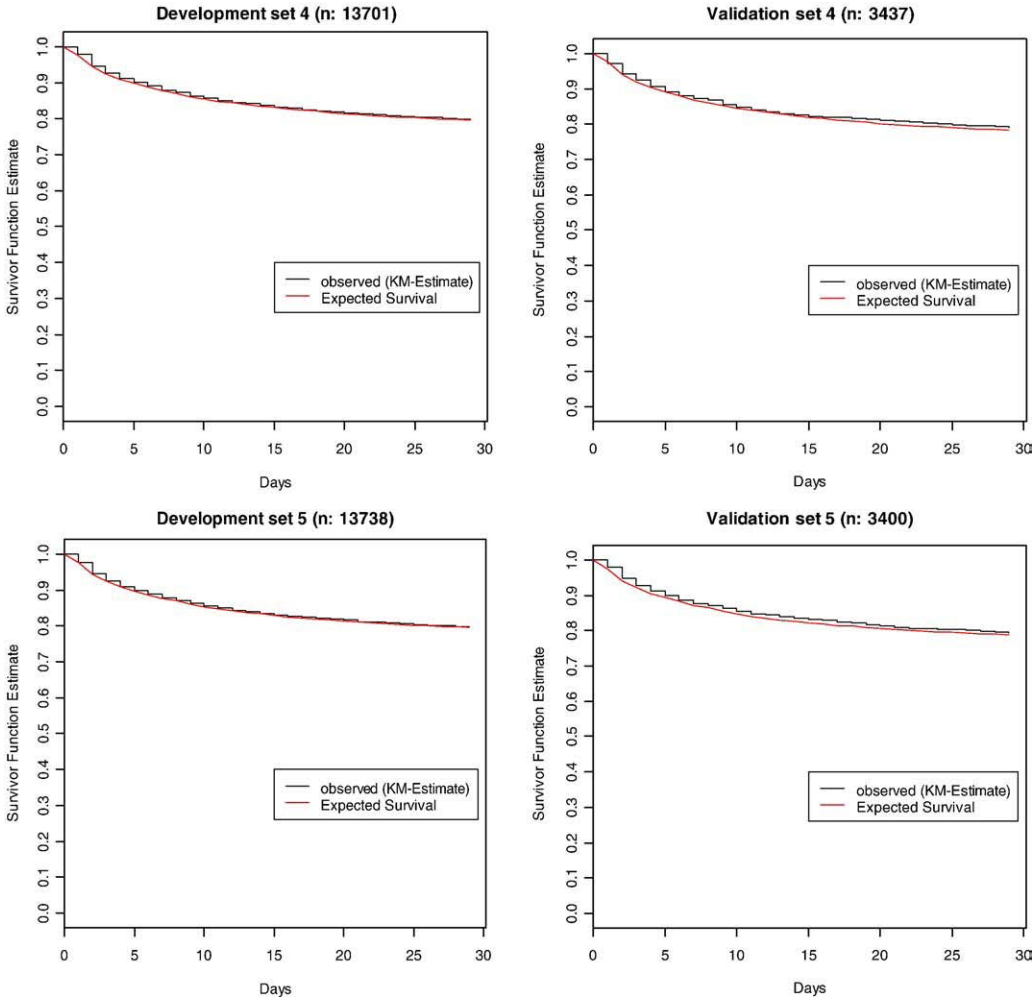

Fig. 2 (*continued*).

patient level) that measure the degree of physiological reserve (age, comorbid diseases), the reasons and circumstances of ICU admission, and the presence and degree of organ dysfunction to model vital status at hospital discharge. Logistic regression is the primary model building technique that is used to predict the outcome of our patients. One of the main shortcomings of this technique, however, is the potential violation of conditional independence in the outcome of interest [9,10] due to the hierarchic nature of the data. Although the logistic regression model is widely used to define factors related to the probability of death, the Cox model is more often used to model the instantaneous risk of nonfatal outcomes.

Actually, the 2 approaches differ in terms of sampling space and entail some simplifications. Thus, the choice of model to use mostly depends on the researcher's intentions. Logistic regression aims at estimating and explaining the probability of the event of interest in a population where all patients are observed over the same period. However, this type of model is frequently used to express the probability of a certain vital status (death) at a variable point in time (date of hospital discharge, which varies among different patients). Therefore, it is usually assumed that length of stay in the ICU does not alter the statistical inference with regard to the probability of death and that time to the event (death) is not important. This assumption is probably not valid; and thus, the use of this method for the specific purpose is questionable.

The Cox model, on the other hand, focuses on modeling a survival function or, conversely, on the distribution of times to the outcome, taking into account the actual individual observation periods. Afterward, this type of model expresses the conditional probability that a subject will develop the outcome of interest per unit of time. The problem of this methodology is an assumption that is frequently not met, namely, that the relative hazards are constant over time—the so-called proportional hazards assumption: that at any time the ratio of the hazard of dying between different risk groups is assumed to be constant. This is certainly not the case in intensive care. Here we have risk groups with high hazards at the beginning, but low hazards if the initial period has been survived, as contrasted to other subgroups in which the hazard may increase with increasing time.

Table 5 Comparison of the survivor rates obtained by logistic regression with random effects, the multilevel Cox proportional hazards model, and Kaplan-Meier estimates in the general population and in major patient subgroups

Subgroup	n	Survivor rate		
		Vital status at 28 d	LR	Cox
Whole data set	17 138	0.80	0.80	0.79
Surgical status				
Nonoperative	8331	0.73	0.73	0.72
Scheduled	5796	0.92	0.92	0.92
Emergency	3011	0.75	0.74	0.75
Trauma	1469	0.82	0.83	0.80
Infection				
No infection	13 663	0.83	0.83	0.83
Community acquired	2302	0.68	0.69	0.68
Hospital acquired	1173	0.59	0.59	0.64
Risk (ICU)				
Low	5665	0.91	0.85	0.85
Middle	5741	0.82	0.81	0.81
High	5732	0.66	0.73	0.73

LR indicates logistic regression.

To deal with this assumption, several methods have been described, such as the use of the subdistribution function (also called the *cumulative incidence function*) as an estimate of the probability of the outcome of interest [7]. This approach was recently followed by Resche-Rigon and coworkers in a small sample of 203 mechanically ventilated patients with reasonable preliminary results [4]. Although some other alternatives have been suggested [11,12] and even preliminarily tested [13,14], no definitive answer for this problem has yet emerged.

Our results provide evidence that logistic regression performs better to forecast vital status at day 28 after ICU admission in this sample. Despite the methodological problems of logistic regression models, as explained above, this seems not so surprising: Cox models are not intended to predict the value of the survival curve at a single time point —in contrast to logistic regression models, which aim to predict one certain value of the survival curve. A further strength of the Cox model is the ability of having the possibility to use information from patients lost to follow-up, censoring them at the point of discharge from ICU, something that cannot be done with logistic regression.

It has to be noted that the selection and weight of the variables estimated by using models of vital status at hospital discharge or in hospital 28 days after ICU admission vital status are different. This has to do with the different sampling intervals and also the different outcome variables studied. This is an important issue because it might explain the often poor performance of outcome prediction models, developed to estimate vital status at hospital discharge, when used to predict vital status at day 28. The PROWESS study for example used the Acute Physiology and Chronic Health Evaluation II scores to control for severity of illness when analyzing in-hospital outcome at 28 days after ICU admission [15]. The Acute Physiology and Chronic Health Evaluation II score, however, was developed to forecast vital status at hospital discharge and was clearly demonstrated not to control adequately for severity of illness 28 days after ICU admission [16]. However, having only 28-day in-hospital mortality (as opposed to true 28-day mortality as used in most randomized controlled trials and by the regulators) is certainly a limitation of this study. The development of a model based on true 28-day mortality thus constitutes an area for future research.

In conclusion, our results demonstrate the need for specific outcome prediction instruments, such as the SAPS 3 28-day score, to estimate in-hospital vital status at 28 days after ICU admission. According to our results, both models (logistic regression or Cox) are suitable; thus, the choice depends on the specific circumstances.

Acknowledgments

This project was supported by a grant from the fund of the Austrian National Bank Project #10995 ONB.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jcrc.2007.11.004.

References

[1] Le Gall 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.
[2] Moreno RP, Metnitz PG, Almeida E, et al. SAPS 3. From evaluation of the patient to evaluation of the intensive care unit. Part 2: development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med 2005;31:1345-55.
[3] Cox DR. Regression models and life-tables (with discussion). J R Stat Soc B 1972;34:107-220.
[4] Resche-Rigon M, Azoulay E, Chevret S. Evaluating mortality in intensive care units: contribution of competing risks analyses. Crit Care 2005;10:R5.
[5] Metnitz PG, Moreno RP, Almeida E, et al. SAPS 3. From evaluation of the patient to evaluation of the intensive care unit. Part 1: objectives, methods and cohort description. Intensive Care Med 2005;31:1336-44.
[6] Hosmer DW, Lemeshow S. Confidence interval estimates of an index of quality performance based on logistic regression estimates. Stat Med 1995;14:2161-72.
[7] Fine J, Gray R. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496-509.
[8] Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
[9] Goldstein H. Multilevel statistical models. London: Arnold; 1995.
[10] Goldstein H, Spiegelhalter DJ. League tables and their limitations: statistical issues in comparisons of institutional performance. J R Stat Soc A 1996;159:385-443.

[11] Gray RJ. Flexible methods for analysing survival data using splines, with applications to breast cancer prognosis. J Am Stat Assoc 1992;87:942-51.
[12] Gray RJ. Spline-based tests in survival analysis. Biometrics 1994;50:640-52.
[13] Kasal J, Jovanovic Z, Clermont G, et al. Comparison of Cox and Gray's survival models in severe sepsis. Crit Care Med 2004;32:700-7.
[14] Nardi A, Schemper M. Comparing Cox and parametric models in clinical studies. Stat Med 2003;22:3597-610.
[15] Bernard GR, Vincent JL, Laterre PF, et al. Efficacy and safety of recombinant human activated protein C for severe sepsis. N Engl J Med 2001;344:699-709.
[16] Carlet J. Prescribing indications based on successful clinical trials in sepsis: a difficult exercise. Crit Care Med 2006;34:525-9.