RIC

ARTÍCULO ORIGINAL

Mortality predictive indexes in non-critical inpatients

Salvador Gómez-García,* Samuel Ponce-de-León,** María V. Hernández-Martínez,** Alejandro E. Macías**,***

* United States-Mexico Border Health Commission, Mexican Section.
** Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán.
*** Escuela de Medicina de la Universidad de Guanajuato.

ABSTRACT

Background. Mortality predictive indexes have not been applied to patients in general wards out of the ICU. **Methods.** Retrolective study aimed to evaluate the value of mortality prediction indexes in a cohort of 944 non-critical patients. Three indexes were evaluated according to their calibration and discriminative power: the Mortality Probability Model II (MPMII), the Simplified Acute Physiology System II (SAPS II) and the Logistic Organ Dysfunction System (LODS). The bivariate calculation of relative risk (RR) to die was performed relative to the group of patients that had an expected probability to die > 10%, calculated by an index. To evaluate the calibration, data were arranged in descending order using the χ^2 goodness-of-fit model. To evaluate discrimination power, ROC curves were used. **Results.** SAPS II, MPM II and LODS predicted significant risks at levels of $P < 0.005$, (RR = 6.56, 4.03 and 3.44, respectively). Regarding the calibration, the null hypothesis was accepted only by using SAPS II ($P = 0.664$). **Conclusions.** The three evaluated indexes each had a good discriminative capacity to detect non-critical inpatients with high risk to die. SAPS II was the best index to predict mortality, as determined by both the bivariate and the calibration analysis. There is no reason for not using mortality predictive indexes for non-critical inpatients.

Key words. Hospital mortality. Forecasting. Inpatients. Quality of health care. Non-critical.

Índices predictivos de mortalidad en pacientes no críticos

RESUMEN

Antecedentes. *Los índices predictivos de mortalidad no se han aplicado rutinariamente a pacientes fuera de las unidades de cuidados intensivos.* ***Métodos.*** *Estudio retrolectivo para determinar el valor de índices predictivos de mortalidad en una cohorte de 944 pacientes no críticos. Se evaluaron tres índices de acuerdo con su calibración y poder discriminador: 1) Mortality Probability Model II (MPMII); 2) Simplified Acute Physiology System II (SAPS II); y 3) Logistic Organ Dysfunction System (LODS). Para cada índice, el cálculo bivariado del riesgo relativo (RR) para morir se efectuó relativo al grupo de pacientes que tenían una probabilidad de morir > 10%. Para evaluar la calibración, los datos se arreglaron en orden descendiente usando el modelo de bondad de ajuste de* χ^2*. Para evaluar el poder de discriminación, se utilizaron curvas ROC.* ***Resultados.*** *Los índices SAPS II, MPM II y LODS predijeron riesgos significativos a niveles de* $P < 0.005$*, (RR = 6.56, 4.03 y 3.44, respectivamente). Respecto de la calibración, la hipótesis nula fue aceptada sólo para el índice SAPS II (*$P = 0.664$*).* ***Conclusiones.*** *Los tres índices evaluados mostraron buena capacidad discriminadora para detectar pacientes no críticos en alto riesgo de muerte, determinado tanto por el análisis bivariado como por su calibración. No existe razón para no utilizar los índices predictivos de muerte en pacientes no críticos.*

Palabras clave. *Mortalidad hospitalaria. Planeación. Calidad en servicios de salud. Pacientes no críticos.*

INTRODUCTION

The quality of hospital attention is a subject which requires evaluation, and we need instruments to compare the results from different hospitals; health institutions can learn a great deal by comparing mutual activities and finding causes for the clinical variations detected.[1-2] Ever since Donabedian created a systematic approach view for the analysis of the quality of medical attention,[3] great advances have been accomplished. One of these is the generation of general indexes to classify the severity of disease and the risk to die, which has simplified decision-making processes and the evaluation of health-care workers' performances.[4,5]

The prognostic indexes can be divided into those for specific diseases and those for general systems. One of the purposes of the scoring systems for intensive care units (ICU) is to obtain data on the likelihood of hospital mortality; with this in mind, experts developed three generations of scales.[6] The first index to appear was the Acute Physiology and Chronic Health (APACHE) index, followed by APACHE II,[7,8] the Mortality Probability Model (MPM I)[9] and the Simplified Acute Physiology System (SAPS I).[10] APACHE II has become the "gold standard" for rating the severity of illness of patients in critical care.[1] Finally, APACHE III, MPM II, SAPS II, and the Logistic Organ Dysfunction System (LODS) constitute the third and last generation of systems based in logistic regression.[6,10-14]

General outcome prediction models using logistic regression analysis have been the cornerstone of performance evaluation in the ICU in recent years. One of the few certain conclusions is that two good models do not always agree in their probability estimates for individual patients, even though they use the same outcome measurements. Unfortunately, unlike many other tools in our clinical armamentarium, there are no standard criteria to compare alternative scoring systems of risk prediction.[15-17]

We have no knowledge of any study in which mortality predictive indexes have been applied to patients in general wards out of the ICU. The objective of the present study was to evaluate the effectiveness of mortality prediction of the MPM II, SAPS II and LODS for non-critical patients.

METHODS

We performed an observational, retrolective study based on a retrospective cohort. The inclusion criteria included all patients over 18 years old discharged from the non-critical wards of the National Institute of Medical Sciences and Nutrition "Salvador Zubirán" (INCNSZ), from January to March 2002. We did not include patients from the respiratory care units, burn and coronary areas, or those in the emergency room. We did not include patients with incomplete medical files, no medical note at admission or discharge, or incomplete laboratory results within 24 hours of admission. The data were obtained from the medical records by a graduate student in the field of clinical epidemiology.

Institution

The INCNSZ is a public tertiary care center in Mexico City. It has a teaching unit, residency programs for most medical specialties and a consolidated program for the prevention and control of hospital infections, which uses active surveillance with adequate laboratory support. Medical attention is provided to more than 200,000 permanent patients with 400 to 450 non-critical hospitalized patients per month. The INCNSZ attends patients from all over Mexico and from all economic and cultural groups. The non-critical hospitalization area is divided in nine wards with 20 beds each.

Analysis of data

The minimum sample size was calculated to be 872 patients, considering the population (n = 200,000), with a 95% confidence interval (CI) and an 80% power. The sampling included the nine hospitalization wards and excluded the emergency room and intensive care.

The bivariate calculation of relative risk to die was performed relative to the group of the patients that had an expected probability to die calculated by an index > 10%, with its respective 95% CI and its hypothesis testing χ^2 MH in the statistical program EPI INFO 96. Mortality prediction models were then evaluated according to their calibration and discriminative power. To evaluate the calibration, data were arranged in descending order using the χ^2 goodness-of-fit model, considering the individual probability to die. Probabilities calculated by the indexes were grouped in an ascending manner into risk percentiles; we calculated the average probability to die per risk percentile and then compared the result with the observed deaths in that risk percentile. Then, they were added and divided by the expected deaths (mean probability to die calculated by the index) as well as by those observed (observed deaths in each risk percentile). These differences were used to calculate a value of χ^2; when the calibration is adequate, the null hypothesis is accepted,[9] meaning that there is not a significant statistical difference between the expected value and the observed one.

To evaluate the discrimination power of each model, ROC curves were used to evaluate the capacity to discriminate those who died against those who did not. Each patient's observed mortality was used to construct a 2 x 2 table.

RESULTS

There were 1,234 discharges during the three months of the study period; we studied the 944 files of patients that met the selection criteria and had medical records containing all of the required data; from them, there were 555 women (58.8%) and 389 men (41.2%). Table 1 shows the descriptive measures of age and days of hospital stay in the sample. Most patients were hospitalized for less than 10 days, and 46 died (4.9%).

Bivariate and multivariate analysis

Taking the probability to die at < 10% calculated by mortality predictive indexes as a cut point for the non-exposed group and at ≥ 10% as post-exposition status, SAPS, MPM and LODS predicted significant risks, as shown in table 2, with the highest risk predicted by SAPS. Table 3 shows the calibration of the three models. As seen, the null hypothesis was accepted only by the use of SAPS II. For MPM II and LODS, the calibration was poor and the null hypothesis was rejected, meaning there were differences between the mortality predicted by the model and the mortality observed.

Table 4 shows that SAPS II is the mortality predictive index with the best discrimination power, inferred by a high value under the ROC curve and a narrower 95% confidence interval.

DISCUSSION

In the present study, all the evaluated indexes had a good discriminative capacity to detect non-critical patients with high risk to die. SAPS II was the best index to predict mortality, determined by both the bivariate and the calibration analysis. In fact, the only model that demonstrated an adequate calibration was SAPS II. On the other hand, LODS and MPM II discriminated correctly, but their calibration was not accurate; therefore, they are not useful for the purpose of predicting mortality in non-critical patients.

Previous studies have evaluated the precision and cost-effectiveness of several indexes for severity of disease and mortality predictors in the ICU (APACHE II, APACHE III, SAPS I, SAPS II and MPM II).[4,18] In a group of 272 admissions, APACHE II proved to be the most precise and the least expensive one.[19]

Beck *et al*. prospectively assessed the impact of low-risk admissions (mortality risk < 10%) on the mortality estimates generated by three prognostic models: APACHE II, APACHE III and SAPS II with a total of 1,497 patients admitted to two general ICU in the United Kingdom. Physiology scores were calculated for the first 24 hours after admission. The three models demonstrated a similar degree of discriminative power, but APACHE III had a lower

Table 1. Age and days of hospital stay in 944 patients: descriptive measures.

Measure	Age	Days of hospital stay
Mean	50.1	9.8
Standard error of mean	0.6	0.2
Median*	50.2	7.8
Standard deviation	18.6	7.3
Rank	79	53
Minimum	18	1
Maximum	97	54
Percentiles*		
25	34.1	5
50	50.2	7.8
75	64.9	12.6

* Calculated from grouped data

Table 2. Bivariate analysis of the probability to die (PTD) calculated by predictive indexes in 944 non-critical inpatients.

Index	Relative risk to die	95% CI	p value (χ^2MH)
SAPS II			
PTD <10	1		
PTD ≥10	6.56	4.8-11.3	< 0.005
MPM II			
PTD <10	1		
PTD ≥10	4.03	2.4-7.6	< 0.005
LODS			
PTD <10	1		
PTD ≥10	3.44	1.8-6.7	< 0.005

Table 3. Calibration of three indexes for mortality prediction in 944 non-critical inpatients, using the goodness-of-fit model.

Percentile	SAPS II O (E)*	MPM II O (E)	LODS O (E)
< 10	0**	1 (1.73)	2 (3.22)
10-20	1 (0.88)**	1 (2.87)	1 (3.22)
21-30	1 (1.45)	2 (3.71)	2 (3.86)
31-40	2 (2.14)	1 (4.72)	1 (4.8)
41-50	4 (2.82)	2 (6.08)	3 (4.8)
51-60	3 (3.98)	5 (7.67)	4 (6.72)
61-70	4 (5.39)	5 (9.92)	5 (10.37)
71-80	4 (7.58)	4 (13.53)	8 (11.36)
81-90	12 (11.4)	5 (19.86)	4 (17.25)
91-100	15 (24.3)	20 (38.26)	16 (27.46)
P value	0.664	< 0 .001	0.005
Restrictions	2	1	3

* O (E) = Observed (expected) deaths.

** Percentiles < 10 and 11-20 analyzed together due to the 0 value in percentile < 10, adding 1 restriction.

Table 4. Discriminative values of three mortality-predictive indexes in 944 non-critical patients.

Model	Area under the curve ROC	95% Confidence interval
SAPS II	0.78	0.71-0.84
MPM II	0.77	0.69-0.84
LODS	0.72	0.64-0.8

ability to correctly discriminate between survivors and non-survivors; this means that its calibration was inferior for both low and high-risk patients. Compared with its predecessor, the APACHE III model assigned a much higher percentage of low-risk patients to the stratum with very low mortality risk (< 2%). The study shows that imperfect characterization of low-risk patients may reduce a model's predictive accuracy for the whole ICU population.[18]

Rowan *et al*. performed a multicenter study to compare the effectiveness of APACHE II and MPM. They found that APACHE II was a better predictor of outcome than MPM when prospectively validated for groups of ICU patients from Britain and Ireland; nevertheless, they concluded that, given the degree of calibration and discrimination of the U.S. APACHE II, it should not be used to predict the outcome of individual patients, but to audit the outcome of intensive care processes.[20]

There is no reason not to use mortality predictive indexes in non-critical inpatients. Our results show that, if properly applied, they may have several potential applications such as: 1) the identification of high-risk patients that may require a different kind of care; 2) a measurement of proficiency between different hospitals; 3) a measurement of proficiency between different areas of the same hospital; and 4) the planning for assigning budgets for general hospitals attending patients of relative high risk.

Our retrolective design obliged us to exclude a number of patients whose records did not have all the required data. We are confident, however, that our sample was objectively evaluated as all of the included patients had complete records. On the other hand, prospective designs may have disadvantages also, such as ethical concerns and bias for providing special care to patients in high risk of dying according to a specific index.

In conclusion, the three predictive indexes evaluated in the present study had statistically significant outcomes and good predictive values. The results may help with the decision-making process during the admission of supposedly non-critical patients. As described, some patients admitted as low-risk are, in fact, at a substantial risk to die. If the SAPS II were calculated routinely for every patient at admission, which is a simple and non-time consuming process, those with a high risk to die would be identified, allowing for their hospitalization in intermediate care units.

REFERENCES

1. Boyd O, Grounds RM. Physiological scoring systems and audit. *Lancet* 1993; 341: 1573-4.
2. Peiró S, Lorenzo S. La difusión a los ciudadanos de los resultados de la asistencia sanitaria. *Rev Calidad Asistencial* 2000; 15: 684-5.
3. Donabedian A. Evaluating the quality of medical care. *Millbank Memorial Fund Quaterly* 1966; 44: 166-206.
4. Cerón DUW, Esponda PJ, Borboya PM, et al. Valor predictivo de los sistemas de calificación de gravedad: comparación de cuatro modelos en tres unidades de terapia intensiva mexicanas incluidas en la base de datos multicéntrica de terapia intensiva. *Revista Mexicana de Medicina Crítica y Terapia Intensiva* 2000; 14: 50-9.
5. Dubois RW, Rogers WH, Moxley JH III, Draper D, Brook RH. Hospital inpatient mortality: is it a predictor of quality? *N Engl J Med* 1987; 317: 1674-80.
6. Lemeshow S, Le Gall JR. Modeling the severity of illness of ICU patients. A system update. *JAMA* 1994; 272: 1049-55.
7. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. APACHE - acute physiology and chronic health evaluation: a physiologically based classification system. *Crit Care Med* 1981; 9: 591-7.
8. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985; 13: 818-29.

9. Le Gall JR, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, et al. A simplified acute physiology score for ICU patients. *Crit Care Med* 1984; 12: 975-7.
10. Lemeshow S, Teres D, Avrunin JS, Gage RW. Refining intensive care unit outcome prediction by using changing probabilities of mortality. *Crit Care Med* 1988; 16: 470-7.
11. 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 hospitalised adults. *CHEST* 1991; 100: 1619-36.
12. 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.
13. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on European /Northamerican multicenter study. *JAMA* 1993; 270: 2957-63.
14. Le Gall JR, Klar J, Lemeshow S, Saulnier F, Alberti C, Artigas A, et al. The logistic organ dysfunction system. A new way to assess organ dysfunction in the intensive care unit. *JAMA* 1996; 276: 802-10.
15. Moreno R, Apolone G, Reis MD. Evaluation of the uniformity of fit of general outcome prediction models. *Int Care Med* 1998; 24: 40-7.
16. Lemeshow S, Klar J, Teres D. Outcome prediction for individual intensive care patients: useful, misused, or abused? *Int Care Med* 1995; 21: 770-6.
17. Angus DC, Pinsky MR. Risk prediction: judging the judges. *Int Care Med* 1997; 23: 363-5.
18. Beck DH, Smith GB, Taylor BL. The impact of low-risk intensive care unit admissions on mortality probabilities by SAPS II, APACHE II and APACHE III. *Anaesthesia* 2002; 57: 21-6.
19. Sánchez VLD. Capacidad discriminativa y costo de los sistemas de calificación de la gravedad de enfermedad en la Unidad de Terapia Intensiva. *Rev Asoc Mex Med Crit Ter Int* 1999; 12: 100-4.
20. Rowan KM, Kerr JH, Major E, McPherson K, Short A, Vessey MP. Intensive Care Society's Acute Physiology and Chronic Health Evaluation (APACHE II) study in Britain and Ireland: A prospective, multicenter; cohort study comparing two methods for predicting outcome for adult intensive care patients. *Crit Care Med* 1994; 22: 1392-1401.

Correspondence and reprint request:

Alejandro E. Macias MD,
Instituto Nacional de Ciencias Médicas
y Nutrición Salvador Zubiran.
Vasco de Quiroga 15,
14000, Mexico, D.F.
Fax : 52 555 5133501
E-mail: aaeemmhh@yahoo.com

Recibido el 5 de octubre de 2006.
Aceptado el 8 de mayo de 2007.