POSSUM scoring system for predicting mortality in surgical patients*

POSSUM ESCORE COMO PREDITOR DE MORTALIDADE EM PACIENTES CIRÚRGICOS

UTILIZACIÓN DEL PUNTAJE POSSUM COMO INDICADOR DE LA MORTALIDAD EN PACIENTES QUIRÚRGICOS

Adriana Cristina Galbiatti Parminondi Elias[1], Tiemi Matsuo[2], Cíntia Magalhães Carvalho Grion[3], Lucienne Tibery Queiroz Cardoso[4], Paulo Henrique Verri[5]

ABSTRACT

This study evaluated the use of the POSSUM (Physiological and Operative Severity Score for Enumeration of Mortality and Morbidity) score for predicting mortality in surgical practice. In this study, 416 surgical patients admitted into ICUs for post-surgical care were analyzed. Both predicted and actual mortality rates were compared, according to four risk groups: 0-4%, 5-14%, 15-49%, 50% and over, and the area under the ROC curve of the POSSUM and APACHE II for mortality. The POSSUM and APACHE II scores overestimated the risk of death. The area under the ROC curve of the POSSUM was 0.762, and under APACHE II was 0.737, suggesting the use of POSSUM as an auxiliary tool to predict the risk of death in surgical patients.

KEY WORDS

Intensive Care Units.
Surgery.
Mortality.
Risk index.

RESUMO

O estudo avaliou a utilização do escore POSSUM (Physiological and Operative Severity Score for Enumeration of Mortality and Morbidity) para predizer a mortalidade na prática cirúrgica.Foram analisados 416 pacientes cirúrgicos com internação na UTI para cuidados de pós-operatório. Foram realizadas comparações entre as taxas de mortalidade predita e observada de acordo com 4 grupos de risco: 0-4%, 5-14%, 15-49%, 50% ou mais, e calculada a área sob a curva ROC do escore POSSUM e APACHE II para a mortalidade. A taxa de mortalidade foi de 22,4%. O escores POSSUM e APACHE II superestimaram o risco de morte, e a área sob a curva ROC do POSSUM foi de 0,762 e a do APACHE II de 0,737, sugerindo a utilização do POSSUM como ferramenta auxiliar na predição de risco de morte em pacientes cirúrgicos.

DESCRITORES

Unidades de Terapia Intensiva.
Cirurgia.
Mortalidade.
Indicador de risco.

RESUMEN

El estudio evaluó la utilización del puntaje POSSUM (Physiological and Operative Severity Score for Enumeration of Mortality and Morbity) para predecir la mortalidad en la práctica quirúrgica. Fueron analizados 416 pacientes quirúrgicos internados en la UTI para cuidados postoperatorios. Fueron realizadas comparaciones entre las tasas de mortalidad estimada y observada, de acuerdo con 4 grupos de riesgo: 0-4%, 5-14%, 15-49%, 50% o más, y calculada el área debajo de la curva ROC del puntaje POSSUM y APACHE II para la mortalidad. La tasa de mortalidad fue de 2,4%. Los puntajes POSSUM y APACHE II superestimaron el riesgo de muerte, y el área debajo de la curva ROC del POSSUM fue de 0,762 y la del APACHE II de 0,737, lo que sugiere la utilización del POSSUM como herramienta auxiliar en la predicción de riesgo de muerte en pacientes quirúrgicos.

DESCRIPTORES

Unidades de Terapia Intensiva.
Cirugía.
Mortalidad.
Indice de riesgo.

* Extracted from the dissertation "Fatores de risco de sepse em pacientes cirúrgicos utilizando os escores APACHE II, SOFA, TISS 28 e POSSUM", State University of Londrina, 2009.

[1] Nurse at the University Hospital Surgery Department at State University of Londrina. Doctoral student at the State University of Londrina. Londrina, PR, Brazil. adrianapaelias@yahoo.com.br [2] PhD in Statistics. Professor at State University of Londrina. Londrina, PR, Brazil. tiemi@uel.br [3] Physician. Master in Internal Medicine. Doctoral student at State University of Londrina. Professor at the University Hospital ICU at State University of Londrina. Londrina, PR, Brazil. cintiagrion@sercomtel.com.br [4] Master in Intensive Care. Doctoral student at State University of Londrina. Chief of the University Hospital Intensive Care Service at State University of Londrina. Londrina, PR, Brazil. lucardoso@sercomtel.com.br [5] Student of the Medicine undergraduate course at State University of Londrina. Holder of a Scientific Initiation grant. Londrina, PR, Brazil. paulohverri@yahoo.com.br

Received: 12/06/2007
Approved: 07/31/2008

Portuguese / English: www.scielo.br/reeusp

INTRODUCTION

Mortality is an important and objective way of measuring results. However, its raw rates are inadequate to define the management of financial resources and monitor the performance of hospital units, as well as to assure quality service. In order to provide quality and safe result measurements, several scoring systems with adjustable risks and stratified for specific populations have been developed to overcome these deficiencies.

Calibrated systems were developed to obtain mortality estimates for septic patient groups[1], oncologic patients[2] and patients submitted to liver transplantations[3]. Among these systems is the Physiological and Operative Severity Score for enUmeration of Mortality and Morbidity (POSSUM), created in 1991 by Copeland and collaborators as a statistical model to predict the surgery risk based on exponential analyses. Through physiological variables of the patient and the surgical procedure, POSSUM assesses the outcomes of the surgical interventions, their complications and the ratio between predicted and observed morbidity and mortality in each death risk range of the population who received this type of care[4-6].

Studies show the efficient results of POSSUM as an appropriate tool to predict mortality in surgery, as well as to compare the performance of different surgery units. Examples include the identification of groups of higher risk patients where quick surgical interventions or better resuscitations in the pre- and intraoperative periods would yield better results, as well as the comparison of performances between different surgery units, where the service that provided care to low-risk patients had lower mortality rates and, consequently, better performance. The calibration of the model is suggested for populations from different countries, due to variations in the nutritional states and mechanisms of defense against disease, as well as diversities in the system of hospital services, which can influence the results of scores developed in a given population and applied to others, questioning the universal use of a single model[7-11].

With the advances in nursing management and the frequent inspections of healthcare auditors, it is necessary to look for surgical patient care assessment systems, not only because they allow for an individualized numeric prediction of mortality, but also because they provide methods that are easily adjusted for changes in the healthcare of the surgery patients in the future. The objective of the study was to evaluate the efficiency of the POSSUM score to predict mortality in surgery patients at a Brazilian teaching hospital.

METHOD

This is a contemporaneous cohort study, with 416 patients aged 14 or older, of both genders, submitted to elective and urgency surgeries, without previous infection, at the surgery center of University Hospital of Londrina, Brazil. The patients were hospitalized at the intensive care unit to receive immediate postoperative care, from January, 2005 to December, 2006.

The nosological file created by the author was filled for all patients in the study, consisting of the following information: name, type of clinic, record number, service number, age, gender, weight, date of admission, surgery preparation (poor, regular, high), type of surgery (elective or urgency), type of anesthesia (general, spinal, peridural, plexus-regional, blocking, sedation), intubation time, surgery time, material sterilization (correct and incorrect), surgery classification according to the criteria of the Hospital Infection Control Committee (clean, potentially contaminated, contaminated, infected), intercurrences in the intraoperative period, number of invasive procedures and the type of surgery performed.

The POSSUM score consists of two parts, which include the physiological variables collected in the preoperative period and the surgical variables collected in the trans- and postoperative period.

The POSSUM score consists of two parts, which include the physiological variables collected in the preoperative period and the surgical variables collected in the trans- and postoperative period. The physiological part of the score includes 12 variables, divided in 4 levels with exponential scores of 1, 2, 4 and 8, represented by the signs and clinical symptoms of each patient, results of biochemical examinations, hematologic investigation and electrocardiographic examinations. In the event in which the variable was not assessed for any reason, the score of 1 was attributed. Some variables were assessed by calculating their averages (clinical signs and symptoms and changes found in the chest radiography). The minimum possible score is 12 and the maximum is 88 points, as recommended by the authors[4-6].

The part of surgery gravity has six variables (magnitude of the surgery, other surgeries within 30 days, blood loss, peritoneal contamination, presence of malignant aspects and surgery type), each of them divided in four levels with scores of 1, 2, 4 and 8.

Data were collected by the study author and entered in an EPI-INFO v. 3.3.2 database by a previously trained intern.

The predicted mortality risk was calculated by the following equation[4]:

$$Ln\left(\frac{R}{1-R}\right) = -7.04 + (0.13*\text{physiological variables}) + (0.16*\text{surgery gravity variables})$$

where R indicates mortality.

The comparative analysis of the averages and medians was performed with Student's t test and Mann-Whitney's test when appropriate. Qualitative variables and outcomes were associated with the chi-square test. Multivariate logistic regression analysis was performed to assess the risk of death of the studied variables. The predictive accuracy for this equation was determined by the ROC (Receiver Operating Characteristic) curve. P values lower than 0.05 were considered significant. The statistical analyses were performed with SAS software (SAS Inst., Inc., Cary, NC, USA).

The study was approved by the Review Board of State University of Londrina (File #026105).

RESULTS

In the 416 analyzed patients, the observed mortality rate was 22.4% (n=93), and the post-surgical infection rate was 39.9% (n=166). Out of the 93 deaths, 72 evolved with post-operatory infection. There was no significant difference in mortality when gender, age, weight, presence of previous diseases and time of pre-surgery hospitalization are considered.

Of the surgeries performed, 64.9% were elective and 35.1% were urgency/emergency. Neurosurgery was the specialty with the greatest number of surgical procedures, and the surgery emergency service *(pronto socorro cirúrgico – PSC)* had the highest mortality rate (Table 1).

Table 1 - Characteristics of the surgery and ICU hospitalization of postoperative patients admitted to the Hospital Universitário de Londrina - Londrina - 2007

Surgery characteristics	Death				P value
	Yes		No		
Type of surgery (n e %)					
Elective	29	31.2	241	74.6	p < 0.001
Urgency/emergency	64	68.8	82	25.4	
Surgical specialties (n e %)					p < 0.001
Cardiac	10	10.8	25	7.7	
Thoracic	1	1.1	34	10.5	
PSC**	36	38.7	29	9.0	
GO***	1	1.1	7	2.2	
Urology	1	1.1	11	3.4	
Neurology	22	23.7	110	34.1	
CAD****	10	10.8	45	13.9	
Vascular	11	11.8	33	10.2	
Ortopedics	1	1.1	29	9.0	
Surgery classification (n e %)					p < 0.001
Clean	56	17.4	267	82.6	
Infected or contaminated*	37	39.78	56	60.22	
Invasive procedures					p < 0.001
Median and interquartile	6.0	5-7	5.0	4-6	
Average and standard deviation	6.12	1.837	5.019	1.742	
Risk levels (n and %)					p < 0.001
Low	30	14.3	182	85.7	
Moderate	36	25.4	104	74.6	
High	27	42.2	37	57.8	
Surgery preparation (n and %)					p < 0.001
Poor	9	26.5	24	73.5	
Regular	53	46.9	60	53.1	
Good	31	11.5	239	88.5	
Days of hospitalization in the PO ICU					p = 0.002
Median and interquartile	5.0	1-10	2.0	1-4	
Average and standard deviation	7.31	8.68	5.19	9.42	
Infection during the PO period (n and %)					p < 0.001
Yes	72	43.4	92	56.6	
No	21	8.4	231	91.6	

* Infected or dirty, contaminated and potentially contaminated. ** Surgery Emergency Service (Pronto Socorro Cirúrgico – PSC). *** Gynecology and Obstetrics. **** Digestive tract surgery (Cirurgia do Aparelho Digestório – CAD).

The patients submitted to clean-rated surgeries had the lowest mortality rate (17.4%) when compared to those submitted to potentially contaminated, contaminated or infected surgeries (39.8%) (p < 0.001)

Patients who did not survive presented the greatest number of invasive procedures, being therefore subject to higher exposure of tissues to external agents (catheters, incisions, drains and others). The surgery risk and quality of surgery preparation levels were significantly different. The time of permanence with the tracheal tube and time of surgery had no significant association with death. However, the median of postoperative hospitalization days at the ICU was related with the outcome – the lower the time necessary for intensive care, the better the prognosis. Death was significantly more frequent in patients with postoperive infection (p < 0.001).

The studied population was distributed in four risk categories, based on values regarding the risk of death calculated by the POSSUM score[12] and APACHE II. The death risk categories (A, B, C and D) were stratified as follows: less than 5%, 5% to 14.99%, 15% to 49.99% and 50% and over. The relative risk of death calculated by POSSUM was higher in the higher risk categories, while the risk of death calculated by APACHE II showed no patients that did not fit categories A and B. There was no progressive relation either between increased risk of death and the risk categories, but when category C was considered as a reference, the risk of death in the higher risk category (D) was still significant (Table 2).

Table 2 - Mortality rates according to the death risk categories calculated by POSSUM and APACHE II for the postoperative patients at the ICU of University Hospital of Londrina - Londrina - 2007

	Death observed in the POSSUM risk categories		Death observed in the APACHE risk categories	
	N (%)	RR (R.I. 95%)	N (%)	RR (R.I. 95%)
A) 0 - 4%	3 (7.1%)	Referência	-	
B) 5 - 14%	13 (9.8%)	1.57 [0.47 - 5.20]	-	
C) 15 - 49%	30 (19.6%)	2.95 [0.95 - 9.20]	54 (14.7%)	Reference
D) 50% or more	47 (52.8%)	7.94 [2.62 - 24.08]	39 (81.3%)	5.54 [4.18 - 7.34]

It was observed that several variables had a significant association with the outcome of death in bivariate analyses (Table 1). However, when analyzed in the logistic regression model, three variables with statistical significance remained: the POSSUM risk categories (0 – 4%, 5 – 14%, 15 – 49%, 50% or more), the level of risk of the surgery (low, moderate and high) and the presence of infection, of any type, during the postoperative period at the ICU (Table 3). The study showed a high rate of infection in the patients analyzed in the postoperative period (39.9%), and this rate increased according to the placement in higher risk categories (P = 0.0001).

There were 93 deaths in the total 416 patients of the study, while the POSSUM score estimated approximately 120 deaths. The ratio between the observed and the expected rates of death (O:E) was 0.77, indicating that the POSSUM scored overestimated the general risk of death and in the higher risk categories (C and D). The APACHE II score also overestimated the general risk of death, but only in the category of risk C (15-49%) (Table 4).

Table 3 – Observed and estimated mortality rates by POSSUM and APACHE. UEL – Londrina – 2007

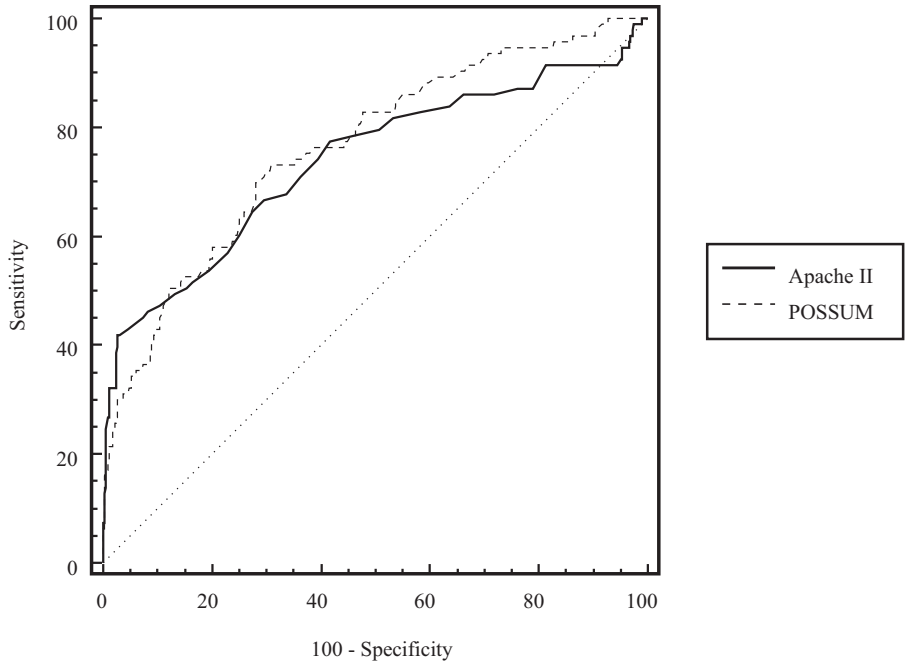
Risk categories	Death				Total
	Estimated		Observed		
	N	%	N	%	
POSSUM					
A (0 - 4%)	1.4	3.3	3	7.1	42
B (5 - 14%)	12.7	9.6	13	9.8	132
C (15 - 49%)	41.2	26.9	30	19.6	153
D (50% or more)	65.1	73.1	47	52.8	89
Total	**120.4**	**28.9**	**93**	**22.4**	**416**
APACHE					
C (15 - 4 9%)	156.3	42.5	54	14.7	368
D (50% or more)	35.1	73.1	39	81.3	48
Total	**191.4**	**46.0**	**93**	**22.4**	**416**

Table 3 - Multivariate analysis of the risk factors for the death outcome in postoperative ICU patients of Hospital Universitário de Londrina, 2005-2006 - Londrina - 2007

Factors	Chance Ratio	R.I. 95%		P value
Risk category - POSSUM				
0 -4%	1.00	Reference		
5 -14%	1.29	0.33	5.00	0.713
15 - 49%	1.72	0.46	6 .43	0.421
50% ou more	7.27	1.95	27.08	0.003
Level of surgical risk				
Low	1.00	Reference		
Average	2.22	1.16	4.25	0.016
High	2.89	1.40	6.00	0.004
Postoperative infection				
No	1.00	Reference		
Yes	6.50	3.64	11.60	<0.001

Analyzing the accuracy of the scores to predict mortality, the POSSUM score, in this study population, had a 0.762 (R.I. 95%: 0.718–0.802) area below the ROC curve, while the area of APACHE II under the ROC curve was 0.737 (R.I. 95%: 0.692–0.779). The difference between both areas was 0.025 and non-significant (p = 0.493) (Figure 1).

Based on the ROC curve, a cut point equal to 0.239 was estimated for the POSSUM score, with 72.3% sensitivity (R.I. 95%: 62.2 – 81.1%) and 69.0% specificity (R.I. 95%: 63.7% – 74.0%). For the APACHE II score, the cut point is 0.482, with a sensitivity of 41.9% (R.I. 95%: 31.8% - 52.6%) and specificity of 97.2% (R.I. %: 94.8% - 98.7%).



DISCUSSION

POSSUM was developed as an auxiliary tool to evaluate the quality of the surgery service[4]. The study showed a mortality rate of 22.5%, considered high when compared with other studies[4, 6-9]. It could be attributed to the fact that only patients submitted to large surgeries and needing postoperative ICU care were included in the study. Some variables are associated to death; factors inherent to the patients, such as the degree of the surgery risk and surgical preparation of the patient prior to the surgery; factors related to the surgery, such as infected or contaminated surgeries, the amount of invasive surgical procedures and factors related to the clinical postoperative evaluation at the ICU and the development of postoperative infection. Knowledge of these risk factors is the basis of service management and indicates which patients are at a higher risk of death and demand more vigilance.

The study population was distributed in four risk categories, based on the reference value of the risk of death

calculated by the POSSUM score. Stratification of the studied patients in relation to these categories of risk of death showed that the relative risk of death increased according to the placement in higher risk categories. This analysis showed the tendency to a death outcome with the increased risk of death calculated by POSSUM with $p < 0.001$, a finding that is similar to other studies[7,12].

The ratio between the observed and estimated mortality (O:E) indicates that the POSSUM score overestimated the risk of death. This result reflects a characteristic of the POSSUM tool which has already been observed by other researchers[7-8,12-13]. These data may also have been influenced by the characteristics of the studied population, which consists of patients who needed ICU hospitalization.

When the POSSUM adjustment was evaluated in relation to the mortality observed in risk brackets, we observed a better distribution of the SMR (Standardized Mortality Rate) – ratio between predicted mortality and observed mortality – only in the lower risk categories (A and B).

Since the prognostic score APACHE II (Acute Physiology and Chronic Health Evaluation) is used at the ICU where the patients received immediate postoperative care, the POSSUM and APACHE II models were compared in order to evaluate the best model to predict risks of death, since APACHE II is not necessarily appropriate for surgery patients, as it does not assess the gravity of the surgical intervention[12].

There were 93 deaths in a total population of 416 patients. The POSSUM score overestimated deaths at 120.42 and the APACHE II score overestimated them at 191.43 deaths. These findings reflect a characteristic of the POSSUM tool which has already been verified by other researchers[8,13]. The tendency to overestimate in higher risk patients is stronger for the POSSUM score when compared with APACHE II.

The Receiver Operating Characteristic (ROC) curve expresses the spectrum of sensitivity and specificity of a given predictor. The POSSUM score, in this study population, had an area under the ROC curve equal to 0.762 (an area equal to 0.5 means a useless predictor, and equal to 1.0 means a perfect predictor). This value is slightly higher than the value found by other authors (0.66[9] to 0.75[10]), occasions when the POSSUM score was considered a good predictor of death. The APACHE II presented, for the same population, an area under the ROC curve equal to 0.737, very similar to the POSSUM score, which was corroborated by the statistical non-significance of the difference between the areas.

Although there have been no Brazilian studies about the utilization of the POSSUM system, nurses, with their theoretical, practical and management reference, could use this tool to help and guide actions to correct and evaluate healthcare for surgery patients.

CONCLUSION

The findings of this study suggest that the POSSUM score can be used as a tool to aid in the prediction of the mortality risk for surgery patients.

REFERENCES

1. Le Gall JR, Lemeshow S, Leleu G, Klar J, Huillard J, Rué M, et al. Customized probability models for early severe sepsis in adult intensive care patients. Intensive Care Unit Scoring Group. JAMA. 1995;273(8):644-50.

2. Staudinger T, Stoiser B, Müllner M, Locker GJ, Laczika K, Knapp S, et al. Outcome and prognostic factors in critically ill cancer patients admitted to the Intensive Care Unit. Crit Care Med. 2000;28(5):1322-8.

3. Angus DC, Clermont G, Kramer DJ, Linde ZWT, Pinsky MR. Short term and long-term outcome prediction with the Acute Physiology and Chronic Health Evaluation II system after orthotropic liver transplantation. Crit Care Med. 2000; 28(1):150-6.

4. Copeland GP, Jones D, Walters M. Possum: a escoring system for surgical audit. Br J Surg. 1991;78(3):355-60.

5. Copeland GP. Surgical scoring, risk assessement and the surgeon. J R Coll Surg Edinb. 1992; 37(3):145-8.

6. Copeland GP. The POSSUM system of surgical audit. Arch Surg. 2002;137(11):15-9.

7. Bollschweiler E, Lubke T, Monig SP, Holscher AH. Evaluation of POSSUM scoring system in patients with gastric cancer undergoing D2 – gastrectomy. BMC Surg. 2005;5:8.

8. Yii MK, Ng KJ. Risk-adjusted surgical audit with the Possum scoring system in a developing country. Br J Surg. 2002; 89 (1):110-3.

9. Brunelli A, Fianchini A, Xiume F, Gesuita R, Mattei A, Carle F. Evaluatin of the POSSUM Scoring system in lung surgery. Thorac Cardiovasc Surgeon. 1998;46(3):141-6.

10. Jones HJS, Cossart L. Risk scoring in surgical patients. Br J Surg. 1999;86(2):149-57.

11. Lubke T, Monig S, Schneider P, Holscher A, Bollschweiler E. Does Charlson – comorbidity index correlate with short-term out come in patients with gastric cancer? Zentralbe Chir. 2003;128(11):970-76.

12. Prytherch DR, Whiteley MS, Higgins B, Weaver PC, Prout WG, Powell SJ. Possum and Portsmouth Possum for predicting mortality. Br J Surg. 1998;85(9):1217-20.

13. Brooks MJ, Sutton R, Sarin S. Comparison of surgical risk score, POSSUM and p-POSSUM in higher-risk surgical patients. Br J Surg. 2005;92(10):1288-92.

Correspondence addressed to: Adriana Cristina Galbiatti P. Elias
Rua Paranaguá, 192 - Apto. 22
CEP 86020-030 - Londrina, PR, Brazil