

# Severity Scores Performance in Neurocritical Care: A Retrospective Cohort Study



**Alexandre Guimarães de Almeida Barros<sup>1</sup>, Guilherme Menezes Mescolotte<sup>2</sup>, Otávio Alves de Souza<sup>3</sup>, Mariana Rabelo de Brito<sup>2</sup>, Natália Lopes Ferreira<sup>1</sup>, Vânia Graner Silva Pinto<sup>1</sup>, Desanka Dragosavac<sup>1</sup> and Antônio Eiras Falcão<sup>1\*</sup>**

<sup>1</sup>Department of Surgery, Faculty of Medical Sciences, State University of Campinas (Unicamp), Brazil

<sup>2</sup>Department of Neurology, Faculty of Medical Sciences, State University of Campinas (Unicamp), Brazil

<sup>3</sup>West Paulista University, Presidente Prudente, Brazil

**Submission:** July 19, 2017; **Published:** August 17, 2017

**\*Corresponding author:** Antônio Eiras Falcão, Department of Surgery, Tessália Viera de Camargo St. 126, University Town Zeferino Vaz, Campinas, São Paulo, Brazil, Zip code 13083-887; Tel: +55(19)3521-9450; Fax: +55(19)3521-8043; Email: aefalcao@gmail.com

## Abstract

Severity scores have wide spread use in intensive care as surrogate markers of disease severity, benchmarking and resource allocation. They are derived from certain cohorts and should go a validation process before their general use in other populations. Nevertheless, they may loss accuracy and prognosis predictability overtime and in different settings. Therefore, constant assess of score's performance is crucial. In this study, we evaluated APACHE II, SAPS 3 and SOFA discrimination and calibration for mortality in a cohort of neurosurgical and neurological patients. Although all have showed good discrimination, APACHE II had a better calibration compared to the others. Hence, our data suggests that APACHE II is a good prognostic model for neurocritical patients.

**Keywords:** Neurocritical care; Prognostic scores; APACHE II; SOFA; SAPS 3; Mortality

## Introduction

Prognostic scores have been widely used in research and in intensive care medicine for disease severity evaluation, benchmarking and optimal resource allocation [1]. These models use certain medical information to predict an outcome chance of happening [2]. Simplified Acute Physiology Score 3 (SAPS 3), Acute Physiology and Chronic Health Disease Classification System II (APACHE II), Sequential Organ Failure Assessment (SOFA) are predictive scores employed to estimate probabilities of death and severity of organ dysfunction in patients admitted in an ICU [3-9]. They are derived from developmental cohorts and may lose precision, after extrapolation to other populations and accuracy overtime [1,10]. Therefore, predictive models should go through external validation and constant update.

Neurocritical care is a subspecialty of critical care medicine that utilize specialized interventions aiming to improve severe neurological and neurosurgical injuries. This may add substantial resource consumption and queries for rational medical practices [11]. Thus, predictive models may play a significant role in this subset of patients.

To assess performance of predictive models in neurocritical care, we performed an external validation analysis using a cohort

of neurological and neurosurgical patients. This is relevant taking into consideration the continuous advancement of care and resource allocation optimization in the field.

## Materials and Methods

### Study design and setting

This was a one center, retrospective cohort study, conducted between January 1, 2013 and December 31, 2016, in the neurological ICU of Unicamp's teaching hospital. This ICU has seven beds and a full time intensive care specialists, nurses, assistants and physiotherapists. Pertinent data on patient's medical history were extracted from an electronic database. All death data corresponded to the ICU stay period. This was an observational study and every clinical decision was at the discretion of the attending physician. Therefore, an informed consent was waived. Local and national ethics committee approved the study.

### Selection of participants

During the study period, all patients with a neurological or neurosurgical diagnosis, aged 18 years or older, requiring an admission to the ICU, were evaluated. Patients with less than 24 hours of ICU stay, who received cardiopulmonary resuscitation before admission or have died during this period were excluded.

**Data gathering**

Data were collected from an electronic database which was fed continuously with patient’s information. This electronic database was protected to hide patient’s identification. Demographic, clinical and laboratory variables of interest were retrospectively collected. All used scores (SAPS 3, APACHE II and SOFA) were automatically calculated with appropriate variables.

**Statistical analysis**

Statistical analyses were performed using MedCalc version 17.6 and Epi Info 7.2.1.0. Continuous variables were presented as mean or median and standard deviation or interquartile range (IQR) as indicated. Categorical variables were presented as absolute values and percentages.

Score discrimination was evaluated by calculating the area under the receiver operating characteristic curve (AUROC) with 95% confidence intervals (CI) and compared with the Hanley-McNeil (1983) test. The discrimination was considered excellent, features are shown in Table 1.

very good, good, moderate and poor with AUROC values of 0.9-0.99, 0.8-0.89, 0.7-0.79, 0.6-0.69 and <0.6, respectively [12]. The correlation between the observed and expected rate of death across all strata of probabilities of death was evaluated by the Hosmer-Lemeshow goodness-of-fit test C-statistic, which assessed the calibration for the test [13,14]. A p value <0.05, two-tailed, and 95% CI were used for logistic regressions.

**Results**

A total of 909 patients fulfilled the inclusion criteria and were enrolled for further analyzes. From those, 611 (67.20%) were hospitalized for neurosurgical procedures and 298 (38,2%) with neurological issues. The mean age was 49.28±16.51, mean ICU length of stay 7,91±10,23 days and mean hospitalization length of stay 22,26±25,37 days. There was a male predominance with 477 cases (52,5%). Observed mortality was 6.5% (59 patients) and mean prognostic scores were 11.13±5.35, 37.51±15.02 and 2.91±2.63 for APACHE II, SAPS 3 and SOFA respectively. Patients

**Table 1:** Population features and outcomes.

Parameters	All Patients	Unit Discharge	
		Dead	Alive
Number or patients	909	59 (06.5%)	850 (93.5%)
Age (years)	49.28±16.51	60.25±14.80	48.52±16.36
Unit Length Stay (days)*	04 (02;09)	07 (04;15)	04 (02;08)
Male	477 (52.5%)	25 (42.4%)	452 (53.2%)
Female	432 (47.5%)	34 (57.6%)	398 (46.8%)
<b>Comorbidities</b>			
Seizures	070 (07.7%)	06(10.2%)	064 (07.3%)
Arterial hypertension	377 (41.5%)	32 (54.2%)	345 (40.6%)
Dyslipidemia	081 (00.9%)	02 (03.4%)	079 (09.3%)
Diabetes Mellitus	152 (16.7%)	15 (25.4%)	137 (16.1%)
COPD	016 (01.8%)	01 (01.7%)	015 (01.8%)
Drugs abuse	030 (03.3%)	01 (01.7%)	029 (03.4%)
Alcoholism	103 (11.3%)	14 (23.7%)	089 (10.5%)
Smoking	214 (23.5%)	16 (27.1%)	198 (21.3%)
Medical Speciality			
Neurosurgery	611 (67.2%)	19 (32.2%)	592 (69.6%)
Neurology	298 (32.8%)	40 (67.8%)	258 (30.4%)
<b>Main Admission Reason</b>			
Tumor Post-Op	299 (32.9%)	04 (06.8%)	295 (34.7%)
Aneurysm Post-Op	102 (11.2%)	08 (13.6%)	094 (11.1%)
Ischemic stroke	089 (09.8%)	16 (27.1%)	073 (08.6%)
Epilepsy Post-Op	084 (09.9%)	00 (-)	084 (09.9%)
Spine Post-Op	082 (09.0%)	01 (01.7%)	081 (09.5%)
Hemorrhagic Stroke	076 (08.4%)	19 (32.2%)	057 (06.7%)
Septic shock	042 (04.6%)	06 (10.2%)	036 (04.2%)
Others	135 (14.9%)	05 (08.5%)	130 (15.3%)

Submitted to Surgery?			
No	253 (27.8%)	28 (47.5%)	225 (26.3%)
Yes, elective	579 (63.7%)	07 (11.9%)	572 (67.3%)
Yes, urgency	077 (8.5%)	24 (40.7%)	053 (6.2%)
Scores			
SAPS 3*	36 (25;47)	56 (52;65)	35 (24;45)
APACHE II*	10 (07;14)	17 (14;21)	10 (07;13)
SOFA*	02 (01;04)	06 (04;08)	02 (01;04)

\*values expressed by median and IQR(Post Op: Postoperative; COPD: Chronic Obstructive Pulmonary Disease; SAPS 3: Simplified Acute Physiology Score 3; APACHE II: Acute Physiology and Chronic Health II)

To assess the performance of prognostic scores in neurocritical patients, we evaluated accuracy of APACHE II, SAPS 3 and SOFA verifying their discrimination and calibration in our cohort of patients. Discrimination was good for all scores. Area under the ROC curves were 0.832 (95%CI: 0.806 to 0.856) for APACHE II, 0.840 (95%CI: 0.814 to 0.863) for SAPS 3, and 0.808 (95%CI: 0.781 to 0.833) for SOFA (Figure 1). Pairwise comparison did not achieve significance according to the Hanley & Mcneil [12] test (APACHE II vs SAPS 3 P: 0,75; APACHE II vs SOFA P: 0,39; SAPS 3 vs SOFA P:0,23). Next, we sought to verify calibration of those scores through the Hosmer–Lemeshow goodness-of-fit C-statistics. Observed p values ascribe a good calibration for APACHE II while SAPS 3 and SOFA had a poor performance (Table 2).

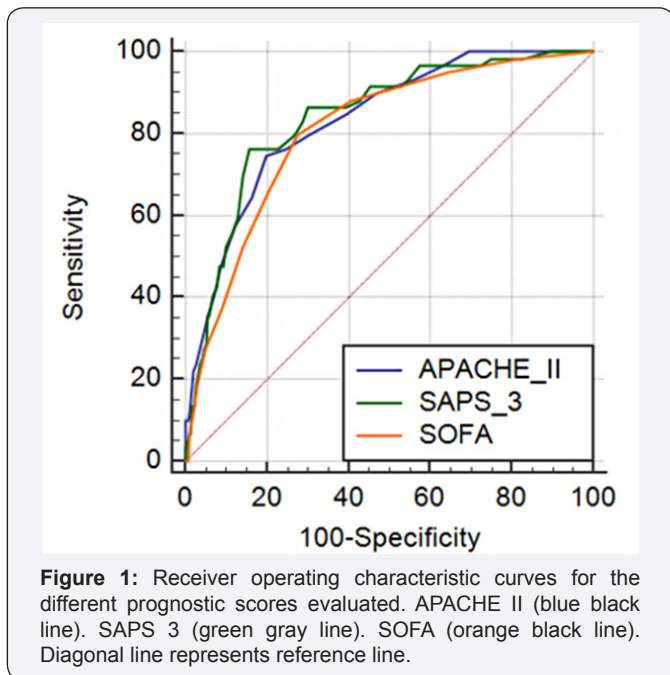


Figure 1: Receiver operating characteristic curves for the different prognostic scores evaluated. APACHE II (blue black line). SAPS 3 (green gray line). SOFA (orange black line). Diagonal line represents reference line.

Table 2: Hosmer-Lemeshow goodness-of-fit C-statistics for the different prognostic scores. A p value more than 0.1 indicates a good calibration.

Score	H-L Statistic	P value
SAPS 3	13.71	0.08
APACHE II	9.22	0.41
SOFA	10.29	0.06

SAPSS: Simplified Acute Physiology Score III; APACHE II: Acute Physiology and Chronic Health Disease Classification System II; SOFA: Sequential Organ Failure Assessment; H-L: Hosmer & Lemeshow

### Discussion

In this study, we evaluated prognostic scores prediction in a cohort of neurological and neurosurgical patients. We observed that although all have a good discrimination, APACHE II are better calibrated than SAPS 3 and SOFA in our settings. Prognostic models are acceptable surrogate indicators for illness severity and organ dysfunction [2,11]. Studies suggest that prediction models must pass through external validation prior of being extrapolated to different populations [1,15]. This is especially important if the new population has features far from score’s developmental cohort. To our knowledge, few studies evaluated prediction models in a subgroup of patients composed of neurological and neurosurgical patients. Based on our data, it is reliable to use APACHE II in neurocritical patients. Moreover, although SAPS 3 and SOFA have good discrimination, both perform poor in subsets of patients as shown by their calibration results. Noteworthy, studies have already noticed poor calibration for prognostic scores in many settings [2,16]. This suggests that even using regional equations, validation must be performed in any new population with adequate sample sizes, if possible, and periodically. Whether a continuous accuracy evaluation or an intermittent process should be performed is unknown.

### Conclusion

Prognostic scores are important markers for stratification and prediction in intensive care medicine. We observed that APACHE II has good discrimination and calibration in neurocritical patients. Our results support its use in this setting of patients. Moreover, Prediction score validation process is crucial for model’s reliability and must be performed in any population before it’s use.

### Acknowledgement

We thank all members of the intensive care medicine and neurology division of Unicamp’s teaching hospital that contributed in this study.

### References

1. Tman DG, Vergouwe Y, Royston P, Moons KGM (2009) Prognosis and prognostic research: validating a prognostic model. *BMJ* 338: b605.
2. Keegan MT, Gajic O, Afessa B (2011) Severity of illness scoring systems in the intensive care unit. *Crit Care Med* 39(1): 163-169.
3. Moreno R, Vincent J, Matos R, Mendonça A, Conraîne F (1999) The use of maximum SOFA score to quantify organ dysfunction / failure in

- intensive care . Results of a prospective, multicentre study. *Intensive Care Med* 25(7): 686-696.
4. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, et al. (1996) The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med* 22(7): 707-710.
  5. Vincent JL, de Mendonça A, Cantraine F, Moreno R, Takala J, et al. (1998) Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units. *Crit Care Med* 26(11): 1793-1800.
  6. Ho KM (2007) Combining Sequential Organ Failure Assessment (SOFA) score with Acute Physiology and Chronic Health Evaluation (APACHE) II score to predict hospital mortality of critically ill patients. *Anaesth Intensive Care* 35(4): 515-521.
  7. Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985) APACHE II: a severity of disease classification system. *Crit Care Med* 13(10): 818-829.
  8. Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, et al. (2005) 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* 31(10): 1345-1355.
  9. Sakr Y, Krauss C, Amaral AC, Réa-Neto A, Specht M, et al. (2008) Comparison of the performance of SAPS II, SAPS 3, APACHE II, and their customized prognostic models in a surgical intensive care unit. *Br J Anaesth* 101(6): 798-803.
  10. Breslow MJ, Badawi O (2012) Severity scoring in the critically ill: part 1--interpretation and accuracy of outcome prediction scoring systems. *Chest* 141(1): 245-252.
  11. Vincent JL, Moreno R (2010) Clinical review: Scoring systems in the critically ill. *Crit Care* 14(2): 207.
  12. Hanley JA, McNeil BJ (1983) A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 148(3): 839-843.
  13. Lemeshow S, Hosmer DWJ (1982) A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 115(1): 92-106.
  14. Paul P, Pennell ML, Lemeshow S (2013) Standardizing the power of the Hosmer-Lemeshow goodness of fit test in large data sets. *Stat Med* 32(1): 67-80.
  15. Nassar AP, Mocelin AO, Nunes AL, Giannini FP, Brauer L, et al. (2012) Caution when using prognostic models: A prospective comparison of 3 recent prognostic models. *J Crit Care* 27(4): 423.
  16. Salluh JIF, Soares M (2014) ICU severity of illness scores. *Curr Opin Crit Care* 20(5): 557-565.



This work is licensed under Creative Commons Attribution 4.0 License  
DOI: [10.19080/OAJNN.2017.5.555662](https://doi.org/10.19080/OAJNN.2017.5.555662)

### Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats  
( Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission

<https://juniperpublishers.com/online-submission.php>