- Melike KORKMAZ TOKER[1]
- Başak ALTIPARMAK[1]
- Canan GÜRSOY[1]
- Ali İhsan UYSAL[1]
- Semra GÜMÜŞ DEMİRBİLEK[1]

CORRESPONDANCE

Melike KORKMAZ TOKER
Muğla Sıtkı Koçman University, Training and Research Hospital, Department of Anesthesiology and Reanimation, Muğla, Turkey

Phone: 02522141323
e-mail: meltoker@gmail.com

[1] Muğla Sıtkı Koçman University, Training and Research Hospital, Department of Anesthesiology and Reanimation, Muğla, Turkey

RESEARCH

ACCURACY AND PERFORMANCE ASSESSMENT OF APACHE IV AND SAPS 3 IN GERIATRIC PATIENTS ADMITTED TO THE INTENSIVE CARE UNIT

ABSTRACT

Introduction: As the world's population is ageing, accurate prognostic prediction in critically ill elderly patients is becoming increasingly important. We aimed to assess the performance of the Acute Physiology and Chronic Health Evaluation IV and Simplified Acute Physiology Score 3 scores in predicting outcome in elderly patients admitted to the intensive care unit.

Materials and Method: In this study, we retrospectively evaluated the patients over 65 years who were followed in the intensive care unit between 2016 and 2018. Acute Physiology and Chronic Health Evaluation IV, Simplified Acute Physiology Score 3 and predicted mortality rate were calculated using a web-based calculator and the predictive mortality performance of these scores were evaluated.

Results: Total mortality was 37% (n=74). Mean Acute Physiology and Chronic Health Evaluation IV and Simplified Acute Physiology Score 3 scores were 87.73±41.24 and 54.87±25.44, respectively. Mean predicted mortality rate according to Acute Physiology and Chronic Health Evaluation IV and Simplified Acute Physiology Score 3 was 41.82±32.76 and 34.60±34.57, respectively. The area under the curve was 0.89 for Acute Physiology and Chronic Health Evaluation IV and 0.91 for Simplified Acute Physiology Score 3. The Hosmer–Lemeshow statistics showed poor calibration for Simplified Acute Physiology Score 3 (p<0.01) and strong calibration for Acute Physiology and Chronic Health Evaluation IV (p> 0.05).

Conclusion: Simplified Acute Physiology Score 3 had better performance in predicting mortality than The Acute Physiology and Chronic Health Evaluation IV in the elderly population.

Keywords: APACHE; Simplified Acute Physiology Score; Critical Care; Mortality

ARAŞTIRMA

YOĞUN BAKIM ÜNİTESİNE BAŞVURAN GERİATRİK HASTALARDA APACHE IV VE SAPS 3'ÜN DOĞRULUK VE PERFORMANS DEĞERLENDİRMESİ

ÖZ

Giriş: Dünya populasyonu yaşlandıkça, yoğun bakımdaki yaşlı hastalarda prognozun doğruluğunu tahmin etmek daha önemli hale gelmektedir. Bu çalışmadaki amacımız, Akut Fizyoloji ve Kronik Sağlık Değerlendirmesi IV ve Basitleştirilmiş Akut Fizyoloji Skoru 3 performansının, yoğun bakım ünitesine yatırılan yaşlı hastalarda sonuçları tahmin etmedeki performansını değerlendirmektir.

Gereç ve Yöntem: Bu çalışmada 2016-2018 yılları arasında yoğun bakım ünitesinde takip edilen 65 yaş üstü hastalar retrospektif olarak inceledi. Akut Fizyoloji ve Kronik Sağlık Değerlendirmesi IV, Basitleştirilmiş Akut Fizyoloji Skoru 3 ve tahmini mortalite oranı web tabanlı bir hesap makinesi kullanılarak hesaplanıp, bu skorların mortaliteyi öngörme performansı değerlendirildi.

Bulgular: Toplam mortalite %37'dir (s=74). Akut Fizyoloji ve Kronik Sağlık Değerlendirmesi IV ve Basitleştirilmiş Akut Fizyoloji Skoru 3 puan ortalaması sırasıyla 87.73±41.24 ve 54.87±25.44 idi. Akut Fizyoloji ve Kronik Sağlık Değerlendirmesi IV ve Basitleştirilmiş Akut Fizyoloji Skoru 3'e göre ortalama tahmini mortalite oranı sırasıyla 41.82±32.76 ve 34.60±34.57 idi. Eğri altındaki alan Akut Fizyoloji ve Kronik Sağlık Değerlendirmesi IV için 0.89 ve Basitleştirilmiş Akut Fizyoloji Skoru 3 için 0.91 idi. Hosmer – Lemeshow istatistikleri ve Basitleştirilmiş Akut Fizyoloji Skoru 3 için zayıf kalibrasyon gösterirken (p<0.01) Akut Fizyoloji ve Kronik Sağlık Değerlendirmesi IV için güçlü kalibrasyon gösterdi (p>0.05).

Sonuç: Basitleştirilmiş Akut Fizyoloji Skoru 3 yaşlı popülasyonda Akut Fizyoloji ve Kronik Sağlık Değerlendirmesi IV'e göre mortaliteyi tahmin etmede daha iyi bir performansa sahiptir.

Anahtar sözcükler: Akut Fizyoloji ve Kronik Sağlık Değerlendirmesi; Basitleştirilmiş Akut Fizyoloji Skoru 3; Kritik bakım; Mortalite

INTRODUCTION

In clinical practice, age is a major barrier to standards of the intensive care unit (ICU). When admitted to ICU, old and very old patients often do not receive adequate diagnostic workup and treatment compared with younger patients. A study on triage decisions in European countries demonstrated that in the fact that elderly patients have more intensive care unit rejections than younger patients and have a higher mortality when admitted, the mortality rate appears reduced for the elderly (1).

The course of management in ICU in addition to the diagnosis and treatment offered must include a prognosis for mortality beginning at admission and during the stay in ICU.

As the world's population is ageing, accurate prognostic prediction in critically ill elderly patients is becoming increasingly important. Several scoring systems have been developed for critically ill patients. Acute Physiology and Chronic Health Evaluation (APACHE) scores and Simplified Acute Physiology Score (SAPS) models are widely used for assessing severity of illness and predicting outcome in critically ill patients. The recent ICU scoring systems SAPS 3 and APACHE IV are powerful revised models (2).

The objective of this study was to assess the performance of APACHE IV and SAPS 3 scores in predicting outcome in a small heterogeneous population of elderly patients admitted to ICU.

MATERIALS AND METHOD

The study was approved by the Mugla Sıtkı Kocman University Medical Faculty institutional review board, and informed consent was obtained from the patients. This retrospective study was conducted in ICU of a 600-bed university teaching hospital. ICU, which had 10 beds, was managed by a closed system. Data from patients aged >65 years old and admitted to ICU between 2016 and 2018 were retrospectively evaluated. Data were collected on patients' age, sex, length of stay in ICU, ICU admission, and outcome. APACHE IV scores were obtained from the worst laboratory findings obtained within 24 h after admission, and SAPS 3 scores were obtained from worst laboratory findings obtained within 1 h after admission. APACHE IV and SAPS 3 scores were calculated using a web-based calculator "http://intensivecarenetwork.com/Calculators/Files/APACHE4.html" and "http://intensivecarenetwork.com/Calculators/Files/SAPS3.html," respectively. Predicted mortality rate (PMR) was calculated using the same web-based calculators. Patients aged <65 years, re-admitted to ICU, with burns, and with insufficient data were excluded. We also excluded patients with an ICU stay<24 h or hospital stay>365 days.

Statistical analysis

The Number Cruncher Statistical System (NCSS) 2007 program (Kaysville, UT, USA) was used for statistical analysis. After evaluation of distribution with Shapiro-Wilk test, normally distributed data and non-normally data according to Shapiro-Wilk test were analyzed with independent t test and the Mann-Whitney U test respectively. Normally distributed data were detailed with mean (SD) and not normally distributed data with median [IQR]. A p value <0.05 was considered statistically significant.

Assessment of the overall accuracy of mortality predictions was performed using the standardized mortality ratio (SMR), and calibration was assessed by the Hosmer–Lemeshow goodness-of-fit C statistic. Discrimination was evaluated using receiver operating characteristic

curves based on calculation of the area under the curve (AUC).

RESULTS

The study was conducted among 200 geriatric patients who were admitted to ICU between January 2016 and February 2018. Of the 200 patients, 24.5% (n=49) were female and 75.5% (n=151) were male. The mean age was 75.22±7.35 years. Comorbidities were detected in 89.5% patients (n=179). Among the patients with comorbidities, 59.0% (n=118) had hypertension, 35.0% (n=70) had diabetes mellitus, 15.5% had chronic obstructive pulmonary disease, 16.5% (n=33) had coronary artery disease and KAH, and 22.5% (n=45) had other diseases. One hundred twenty patients (60%) were admitted to ICU from the operating theater, 52 patients (26%) from the emergency department, and 28 patients (14%) from the wards. Among the 120 postoperative patients, 38.5% (n=77) had undergone emergency surgery and 21.5% (n=43) had undergone elective surgery. Total mortality rate was 37% (n=74) WW(Table 1).

The mean APACHE IV and SAPS 3 scores were 87.73±41.24 and 54.87±25.44, respectively. The mean PMR was 41.82±32.76 according to APACHE IV and 34.60±34.57 according to SAPS 3 (Table 2).

Table 3 shows actual mortality and predicted mortality from the scoring systems. Mortality was correctly predicted with the SAPS 3 model.

The SAPS 3 and APACHE IV scores were significantly higher in patients who died (p<0.001). Predicted mortality rate according to SAPS 3 and APACHE IV scores was also significantly higher in patients who died (p<0.001).

Standardized mortality rate was 0.900 (95% CI, 0.713–1.127) according to APACHE IV scores and 1.042 (95% CI, 0.824–1.301) according to SAPS 3 scores. The Hosmer–Lemeshow statistic showed poor calibration for SAPS 3 (p<0.01) but strong calibration for APACHE IV (p>0.05) (Table 4).

There was a statistically significant difference between SAPS 3 and APACHE IV scores according to mortality (p<0.001). We, therefore, decided to calculate cut-off points for APACHE IV and SAPS 3 scores according to mortality by receiver operating characteristic analysis and diagnostic scanning tests. Results for APACHE IV score were cut-off point of 101, sensitivity of 86.49%, specificity of 80.95%, positive predictive value of 72.73, and negative predictive value of 91.07. Results for SAPS 3 score were cut-off point of 53, sensitivity of 89.19%, specificity of 76.19%, positive predictive value of 68.75, and negative predictive value of 92.31 (Table 5).

Area under curve was 88.7% with 2.4% standard deviation for APACHE IV and 90.7% with 2% standard deviation for SAPS 3.

The predicted level of mortality for SAPS 3 scores was higher than that for APACHE IV score, but the difference was not statistically significant (p=0.280 and p>0.05, respectively).

Table 1. Patient demographics and characteristics.

Descriptive characteristics		n (%)
Age (year)	Min–Max (Median)	65–92 (75)
	Mean±SD	75.22±7.35
Gender	Female	49 (24.5)
	Male	151 (75.5)
Main comorbidities*	No	21 (10.5)
	Yes	179 (89.5)
	HT	118 (59.0)
	DM	70 (35.0)
	COPD	31 (15.5)
	CAD	33 (16.5)
	Other	45 (22.5)
Route of admission	Emergency room	52 (26.0)
	OR/recovery room	120 (60.0)
	Ward	28 (14.0)
Admission type	No surgery	80 (40.0)
	Emergency surgery	77 (38.5)
	Elective surgery	43 (21.5)
Length of stay at ICU	Min–Max (Median)	1–78 (4)
	Mean±sd	7.88±13.92
Patient prognosis	Transferred to ward	126 (63.0)
	Exitus	74 (37.0)
*Multiple additional disease choices have been made HT: Hypertension, DM: Diabetes Mellitus, COPD: Chronic Obstructive Pulmonary Disease, CAD: Coronary Artery Disease, OR: Operating Room		

Table 2. Distribution according to APACHE IV scores and SAPS 3.

	Min–Max (Median)	Mean±sd
APACHE IV	21–173 (77)	87.73±41.24
APACHE IV PMR	0.7–94.1 (35.6)	41.82±32.76
APACHE IV LOS	0.9–12.3 (7.3)	7.44±2.34
SAPS 3	13–92 (51)	54.87±25.44
SAPS 3 PMR	0–86 (17)	34.60±34.57
APACHE: Acute Physiology and Chronic Health Evaluation, SAPS: Simplified Acute Physiology Score, PMR: Predicted Mortality Rate, LOS: Length of stay		

Table 3. ICU mortalities predicted by the two models.

Models	Actual Mortality	Predicted Mortality
APACHE IV	0.37	0.42
SAPS 3	0.37	0.35
APACHE: Acute Physiology and Chronic Health Evaluation, SAPS: Simplified Acute Physiology Score		

Table 4. Hosmer–Lemeshow statistic results for APACHE IV and SAPS 3 according to mortality.

	H-L statistics	p	SMR (95% CI)
APACHE IV	7.981	**0.334**	0.900 (0.713–1.127)
SAPS 3	25.254	**0.001****	1.042 (0.824–1.301)
APACHE: Acute Physiology and Chronic Health Evaluation, SAPS: Simplified Acute Physiology Score, SMR: Standardized mortality rate			

Table 5. Diagnostic scanning tests and receiver operating characteristic curve results for APACHE IV scores and SAPS 3.

	Diagnostic Scan					ROC Curve		p
	Cut-off	Sensitivity	Specifity	Positive Predictive Value	Negative Predictive Value	Area	95% Confidence Interval	
APACHE IV	**≥101**	86.49	80.95	72.73	91.07	**0.887**	0.839–0.935	**<0.001****
SAPS 3	**≥53**	89.19	76.19	68.75	92.31	**0.907**	0.867–0.946	**<0.001****
APACHE: Acute Physiology and Chronic Health Evaluation, SAPS: Simplified Acute Physiology Score								

DISCUSSION

Our study demonstrated that the APACHE IV model showed strong calibration whereas the SAPS 3 model showed poor calibration. Both models showed good discrimination. The APACHE IV model overestimated the observed mortality rate.

Results for the SAPS 3 model in our study were not totally unanticipated as many studies have reported results similar to those of as our study (3–5). Nassar

et al. determined the performance of APACHE IV, SAPS 3, and Mortality Probability Model III (MPM0-III) in patients at three medical surgical ICUs in Brazil. In their study, all three models had poor calibration but very good discrimination (3). Although we found the same results, the patients were completely different from that in the study by Nassar et al. in which the three diagnostic models were assessed. Our study population mainly comprised surgical patients. Sixty percent were surgical patients admitted to ICU after

a surgery, and an additional 21.5% were patients undergoing elective surgery. Although postsurgical patients were generally at low risk, they admitted to ICU because of the absence of post anesthesia care unit in our hospital.

External validation is crucial for achieving definite evaluation of these prognostic models (6). To the best of our knowledge, this is the first study to validate the APACHE IV model and compare it with the SAPS 3 model among geriatric patients in ICU in Turkey. In a Dutch study, the APACHE IV model showed very good discrimination (AUC, 0.87) but poor calibration (Hosmer–Lemeshow statistic, 822.67) (7); in a US study, APACHE IV again showed very good discrimination (AUC, 0.86) (8). Our results are consistent with those of studies showing good discriminatory power of the APACHE IV model (AUC, 0.88). However, in our study, the APACHE IV model showed strong calibration (Hosmer–Lemeshow statistic, 7.981). The main goal in applying a scoring system at admission to ICU is to predict mortality of the patient and determine management; hence, it is always required to calibrate the scoring system before generating it conventional (9). Thus, strong calibration of the APACHE IV score in our study makes this score acceptable for geriatric patients admitted in our ICU.

The SAPS 3 admission score is based on records obtained within the 1st h after acceptance of a geriatric patient to ICU (2). Almost half of the predictive power of original SAPS 3 was generated by gathering trauma patients' data prior admission to ICU. Scores recorded after the first 24 h following ICU admission reflected standard care rather than actual clinical status. Because of this major advantage of SAPS 3, the predicted level of mortality of SAPS 3 was higher than that of APACHE IV scores. We observed a poor calibration for SAPS 3 (Hosmer–Lemeshow statistic, 25.254), whereas discrimination was very good (AUC, 0.91). This model was globally evaluated in ICUs; results showed poor calibration but good discrimination (10–14). These results can be explained by the higher proportion of surgical patients with worsened medical conditions. It is important to remember that surgical patients have different physiological and functional characteristics from medical patients and these may influence the prognosis (15).

Keegan et al. concluded from their study that the APACHE IV model had better discriminatory capability than the SAPS 3 model (8). On the other hand, studies in patients with acute kidney injury and acute coronary artery syndrome showed that the two models had similar discriminatory performance (10,16). Poor model calibration can also be correlated with a higher proportion of patients who are at low risk (17). These differences can be explained by discrepancies in study populations and territorial variation in do not resuscitate decisions.

Although the SAPS 3 model has better discriminative power and a tendency to estimate mortality more accurately, its lack in calibration makes it less suitable than the APACHE IV model. Aggarwal et al. suggested that lack of acceptable calibration, regardless of good discrimination power, should result in rejection of a scoring system (18).

The results of our survey should be carefully interpreted because of present limitations. Our study was conducted at a single-center, mixed surgical and medical ICU, which limits the capacity of our results to be generalized to other ICUs. Another limitation is associated with its retrospective design. Furthermore, customization might have supplied a better calibration for the SAPS 3 model, but we decided not to try this path because our objective was to evaluate original scores.

In conclusion, our study showed that APACHE IV and SAPS 3 models had very good discriminative power for predicting mortality in geriatric patients admitted to ICU. Although the SAPS 3 model had poor calibration, it had better performance in predicting mortality than the APACHE IV model.

REFERENCES

1. Sprung CL, Artigas A, Kesecioglu J, et al. The Eldicus prospective, observational study of triage decision making in European intensive care units. Part II: intensive care benefit for the elderly. Crit Care Med 2012;40(1):132-8. (PMID:22001580).

2. Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. Critical Care 2010;14(2):207-9. (PMID:20392287).

3. Nassar AP Jr, Mocelin AO, Nunes ALB, et al. Caution when using prognostic models: A prospective comparison of 3 recent prognostic models. J Crit Care. 2012;27(4):423.e1-7. (PMID:22033059).

4. Beck DH, Smith GB, Pappachan JV, Millar B. External validation of the SAPS II, APACHE II and APACHE III prognostic models in South England: A multicentre study. Intensive Care Med 2003;29(2):249-56. (PMID:12536271).

5. Pappachan JV, Millar B, Bennett ED, Smith GB. Comparison of outcome from intensive care admission after adjustment for case mix by the APACHE III prognostic system. Chest 1999;115(3):802-10. (PMID:10084495).

6. Terrin N, Schmid CH, Griffith JL, D'Agostino RB, Selker HP. External validity of predictive models: A comparison of logistic regression, classification trees, and neural networks. J Clin Epidemiol 2003;56(8):721-9. (PMID:12954463).

7. Brinkman S, Bakhshi-Raiez F, Abu-Hanna A, et al. External validation of Acute Physiology and Chronic Health Evaluation IV in Dutch intensive care units and comparison with Acute Physiology and Chronic Health Evaluation II and Simplified Acute Physiology Score II. J Crit Care 2011;26(1):105.e11-8. (PMID:20869840).

8. Keegan MT, Gajic O, Afessa B. Comparison of APACHE III, APACHE IV, SAPS 3, and MPM0III and influence of resuscitation status on model performance. Chest 2012;142(4):851-8. (PMID:22499827).

9. Fadaizadeh L, Tamadon R, Saeedfar K, Jamaati HR. Performance assessment of Acute Physiology and Chronic Health Evaluation II and Simplified Acute Physiology Score II in a referral respiratory intensive care unit in Iran. Acta Anaesthesiol Taiwanica 2012;50(2):59-62. (PMID:22769859).

10. Silva VTCE, De Castro I, Liaño F, et al. Performance of the third-generation models of severity scoring systems (APACHE IV, SAPS 3 and MPM-III) in acute kidney injury critically ill patients. Nephrol Dial Transplant 2011;26(12):3894-901. (PMID:21505093).

11. Soares M, Silva UVA, Teles JMM, et al. Validation of four prognostic scores in patients with cancer admitted to Brazilian intensive care units: results from a prospective multicenter study. Intensive Care Med 2010;36(7):1188-95. (PMID:20221751).

12. Lee H, Shon YJ, Kim H, Paik H, Park HP. Validation of the APACHE IV model and its comparison with the APACHE II, SAPS 3, And Korean SAPS 3 models for the prediction of hospital mortality in a Korean surgical intensive care unit. Korean J Anesthesiol 2014;67(2):115-22. (PMID:25237448).

13. Khwannimit B, Bhurayanontachai R. The performance and customization of SAPS 3 admission score in a Thai medical intensive care unit. Intensive Care Med 2010;36(2):342-6. (PMID:19756506).

14. Sakr Y, Krauss C, Amaral ACKB, et al. 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 2008;101(6):798-803. (PMID:18845649).

15. Sanchez-Hurtado LA, Angeles-Velez A, Tejeda-Huezo BC, Garcia-Cruz JC, Juarez-Cedillo T. Validation of a prognostic score for mortality in elderly patients admitted to Intensive Care Unit. Indian J Crit Care Med 2016;20(12):695-700. (PMID:28149026).

16. Nassar Junior AP, Mocelin AO, Andrade FM, et al. SAPS 3, APACHE IV or GRACE: Which score to choose for acute coronary syndrome patients in intensive care units? Sao Paulo Med J 2013;131(3):173-8. (PMID:23903266).

17. 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(1):21-6. (PMID:11843737).

18. Aggarwal AN, Sarkar P, Gupta D, Jindal SK. Performance of standard severity scoring systems for outcome prediction in patients admitted to a respiratory intensive care unit in North India. Respirology 2006;11(2):196-204. (PMID:16548906).