original Article

J Nepal Health Res Counc 2026 Jan-March; 24 (70): 151-158

# Predictive Performance of the Acute Physiology and Chronic Health Evaluation II, Simplified Acute Physiology Score II and Sequential Organ Failure Assessment score in Critically ill Patients in Intensive Care Unit

Raj Kumar Mehta,[1] Hafizah Che Hassan,[1] Bishnu Bista,[2] Mamata Sharma Neupane[2]

[1]Lincoln University College, Petaling Jaya, Malaysia, [2]Chitwan Medical College, Bharatpur, Nepal.

## ABSTRACT

**Background:** Severity scoring systems are widely used in intensive care units to assess illness severity, predict mortality, and support clinical decision-making. However, most systems, including the Acute Physiology and Chronic Health Evaluation II, Simplified Acute Physiology Score II, and Sequential Organ Failure Assessment, were developed in high-income countries, and their validity in low- and middle-income settings such as Nepal remains uncertain. This study evaluated and compared the predictive performance of Acute Physiology and Chronic Health Evaluation II, Simplified Acute Physiology Score II, and Sequential Organ Failure Assessment among critically ill patients admitted to a tertiary hospital intensive care units in Nepal.

**Methods:** A prospective observational study was conducted among 501 adult intensive care units' patients over one year. Acute Physiology and Chronic Health Evaluation II, Simplified Acute Physiology Score II, and Sequential Organ Failure Assessment scores were calculated using the worst physiological values within the first 24 hours of intensive care units' admission. Patients were followed until intensive care units discharge and categorized as survivors or non-survivors. Discrimination was assessed using the area under the receiver operating characteristic curve, and calibration was evaluated using the Hosmer–Lemeshow test and standardized mortality ratio. The Mann–Whitney U test compared median scores between outcome groups.

**Results:** The intensive care units mortality rate was 36.9%. Sequential Organ Failure Assessment demonstrated excellent discrimination (Area under receiver operating characteristics curve 0.895), outperforming Simplified Acute Physiology Score II (0.748) and Acute Physiology and Chronic Health Evaluation II (0.722). Sequential Organ Failure Assessment showed good model fit ($p=0.129$), while Acute Physiology and Chronic Health Evaluation II and Simplified Acute Physiology Score II had poor calibration by Hosmer–Lemeshow testing. Although Acute Physiology and Chronic Health Evaluation II and Simplified Acute Physiology Score II showed moderate standardized mortality ratio calibration, SOFA significantly overestimated mortality (Standardized mortality ratio 0.38, $p<0.001$). All three scores were significantly higher among non-survivors ($p<0.001$).

**Conclusions:** All three scoring systems effectively differentiated survivors from non-survivors. Sequential Organ Failure Assessment demonstrated superior discriminative ability and overall predictive performance, making it the most reliable mortality predictor in this intensive care units setting, although recalibration may be required for accurate absolute mortality estimation in resource-limited contexts.

**Keywords:** APACHE II score; critical illness; intensive care units; SAPS II score; SOFA.

## INTRODUCTION

Critical care focuses on managing patients with life-threatening illnesses requiring advanced and organized support in intensive care units (ICUs).[1] Accurate assessment of illness severity is essential for guiding treatment decisions, allocating resources, benchmarking performance, and predicting outcomes.[2,3] Several prognostic models have been developed to objectively quantify physiological derangement and estimate mortality risk.[4] Widely used systems include the Acute Physiology and Chronic Health Evaluation (APACHE) models, [5-8] the Simplified Acute Physiology Score (SAPS) models,[9-11] and the Sequential Organ Failure Assessment (SOFA) score.[12] Among these, APACHE II, SAPS II, and SOFA are commonly applied in multidisciplinary ICUs. However, these tools were primarily developed and validated in high-income countries, and their predictive

**Correspondence:** Raj Kumar Mehta, Lincoln University College, Malaysia. Email: mehtarajkraj@gmail.com, Phone: +9779852024779.

accuracy may vary in low- and middle-income settings. Evidence from Nepal remains limited, and available studies are relatively few, highlighting the need to further evaluate the performance and applicability of these scoring systems in local ICU populations.

## METHODS

A prospective observational design was used to examine the predictive performance of Severity Scoring Systems on the outcome of Critically Ill patient in ICU. This study was conducted at adult intensive care units (ICUs). The population of this study was all the critically ill patients admitted in critical care units of Chitwan Medical College Teaching Hospital. Required sample size for this study was calculated by using Sample size for studies with binary test outcome i.e., Sample size for adequate sensitivity[13] and calculated sample size was 501. Non probability, consecutive sampling technique was used for selection of critically ill patients. The study included critically ill patients who stayed in ICU at least 24 hours after enrolment and with more than 18 years of age and willing to participate voluntarily were included in this study. Burn injury patients, patients who died within first 24 hours of admission to ICU after enrolment, those patients who were left ICU against medical advice were excluded from this study.

The study was carried out over one year period from March 2022 to February 2023. Data were collected through use of a validated standardized tool, APACHE II, SAPS II and SOFA scoring system scale. APACHE II was defined as a severity scoring system used to estimate the risk of mortality among critically ill patients. In this study, the APACHE II score was calculated using the worst physiological values recorded within the first 24 hours of ICU admission, including 12 physiological variables, age points, and chronic health status, with total scores ranging from 0 to 71. Higher scores indicate greater disease severity and higher risk of mortality. SAPS II was defined as a prognostic scoring system used to estimate hospital mortality among ICU patients. In this study, SAPS II was calculated using 17 variables comprising physiological parameters, age, type of admission, and underlying disease conditions recorded during the first 24 hours of ICU admission. The total score ranges from 0 to 163, with higher scores indicating increased severity of illness and mortality risk. SOFA was defined as a scoring system used to assess the degree of organ dysfunction in critically ill patients. In this study, SOFA score was calculated based on six organ systems (respiratory, cardiovascular, hepatic, coagulation, renal, and neurological) using the worst values recorded within the first 24 hours of ICU admission. Each organ system is scored from 0 to 4, with a total score ranging from 0 to 24. Higher scores indicate greater organ dysfunction and increased risk of mortality. A structured questionnaire was used for collecting information about demographic variables and clinical characteristics. The study was carried out after the approval of research ethics committee of university and Nepal Health Research Council (Ref: LUC/REC/INT/111/2022 and NHRC/071/2022) and permission for data collection was granted by the study site (Ref: CMC-IRC/078/079-150). The informed consent was obtained from each respondent or legally authorized representative (LAR). Strict confidentiality of information was maintained. APACHE II, SAPS II and SOFA score points were assigned to all patients by calculating their individual score and predicted risk of mortality was also calculated. The scores were computed using the most unfavourable physiological data obtained within the initial 24 hours following ICU admission. The patients were followed up till discharge from ICUs. The final outcome was labelled as alive or dead at the time of discharge. All data were entered into the Statistical Product and Service Solutions (IBM-SPSS) 25.0 version software for statistical analysis.

Data were summarized by using descriptive and inferential statistical methods. Data normality was assessed with Kolmogorov-Smirnov (K-S) and Shapiro-Wilk (S-W) test and data was not normally distributed (K-S and S-W for APACHE II, p=0.004 and p=0.000; SAPS II, p=0.000 and p=0.000; SOFA, p=0.000 and p=0.000). The patient characteristics and score of all the severity scoring system was summarized using median and interquartile range (IQR) for continuous variables, frequency and percentage for categorical variables. Performance of models were assessed by discrimination and calibration.

Accuracy and discrimination of the severity scoring system model was evaluated by calculating area under a receiver operating characteristic (AUROC) curve with a 95% confidence interval (CI), the optimal cutoff for predicting mortality was defined as that with the highest value of Youden's J index (sensitivity+ specificity−1), and all AUROC comparisons were performed using DeLong's method. Furthermore, AUROCs of above 0.90, 0.80-0.89, and 0.70-0.79 are considered excellent, good and moderate respectively. The calibration of the severity scoring system model was evaluated by Hosmer-Lemeshow test for the goodness-of-fit and Standardized mortality ratio (SMR). The Hosmer–Lemeshow test is a statistical test for goodness of fit and calibration for logistic regression models. A large value of Chi-squared (with small p-value < 0.05) indicates poor fit and small Chi-squared values (with larger p-value closer to 1) indicate a good logistic regression model fit. The calibration of the severity scoring system model was also analysed by calculating Standardized mortality ratio (SMR and 95% CI) with OpenEPI calculator. The SMR referred to the ratio of actual number of deaths and number of deaths that the severity of score predicted. A ratio greater than 1.0 indicates that more mortality has occurred than would have been expected, a value of 1.0 is considered normal or as expected, while a ratio less than 1.0 indicates that less mortality has occurred i.e., better than expected. The severity system score was compared with the final outcome. The Mann-Whitney U test was used to compare the median APACHE II, SAPS II, and SOFA scores between survivors and non-survivors. $p<0.05$ was considered statistically significant.

## RESULTS

A total of 501 critically ill patients admitted in adult ICUs were enrolled. The age of the patients ranged between 18 to 86 years. The median age of the critically ill patient was 48.96 (IQR:65-30). Males comprised 306 (61.1%) and female comprised 195 (38.9%) of the total number of patients. Of the 501 patients, 185 (36.9%) died and 316 (63.1%) survived. The median severity scoring system scores for critically ill patients was 22 (IQR: 27-17, min. 8 and max. 40) for APACHE II, 48 (IQR: 57-36.50; min. 20 and max. 102) for SAPS II and 13 (IQR: 15-11; min. 5 and max. 23) for SOFA respectively.

The ROC curve analysis assesses the predictive performance of three severity scoring systems for patient outcomes in critically ill ICU patients (Table 1 and Figure 1).

**Table 1. Receiver Operating Characteristic curve (ROC) analysis for the APACHE II, SAPS II and SOFA Scoring System for outcome prediction. n=501**

| Scoring Systems | Cut off point | AUROC | p value | Sensitivity | Specificity | 95% CI | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower | Upper |
| APACHE II | 22.50 | 0.722 | <0.001 | 0.654% | 0.652% | .676 | .767 |
| SAPS II | 48.50 | 0.748 | <0.001 | 0.714% | 0.649% | .705 | .790 |
| SOFA | 13.50 | 0.895 | <0.001 | 0.892 | 0.763% | .868 | .922 |

APACHE II: Acute physiology and chronic health evaluation II; SAPS II: Simplified acute physiology score II; SOFA: sequential organ failure assessment score, and AUROC: area under the receiver operating characteristics curve

APACHE II has the moderate discrimination ability (AUROC=0.722; 95% CI: 0.676 -0.767). SAPS II have the better discrimination ability than APACHE II but still moderate (AUROC=0.748; 95% CI: 0.705 -0.790) and SOFA has the excellent discrimination ability (AUROC=0.895; 95% CI: 0.868 -0.922). SOFA has the highest AUROC, indicating it is the most accurate predictor of patient outcomes among the three scoring systems. Thresholds of three scoring systems (Cut-off points, APACHE II: 22.50; SAPS II: 48.50; SOFA: 13.50) represent the scores that best differentiate between survival and mortality in ICU patients. SOFA has sensitivity 0.892, Specificity 0.763 indicating highly accurate in detecting true positives (high sensitivity) while maintaining good specificity. SAPS II have sensitivity 0.714, specificity 0.649 indicating moderate accuracy in predicting outcomes. APACHE II has sensitivity 0.654, specificity 0.652 indicating least accurate of the three models. SOFA has the best balance of sensitivity and specificity, meaning it is the most reliable predictor. SOFA is the best predictive model for ICU patient outcomes, given its highest AUROC (0.895) and best sensitivity-specificity balance. SAPS II perform moderately well but is less reliable than SOFA. APACHE II has the lowest predictive power among the three models. The Hosmer-Lemeshow test assesses the goodness-of-fit for logistic regression models used in prediction. A p-value > 0.05 indicates a good fit, while a p-value < 0.05 suggests a poor fit. SOFA provides a better model fit (p=0.129) compared to APACHE II (p=0.016) and SAPS II (p=0.004) (Table 2).

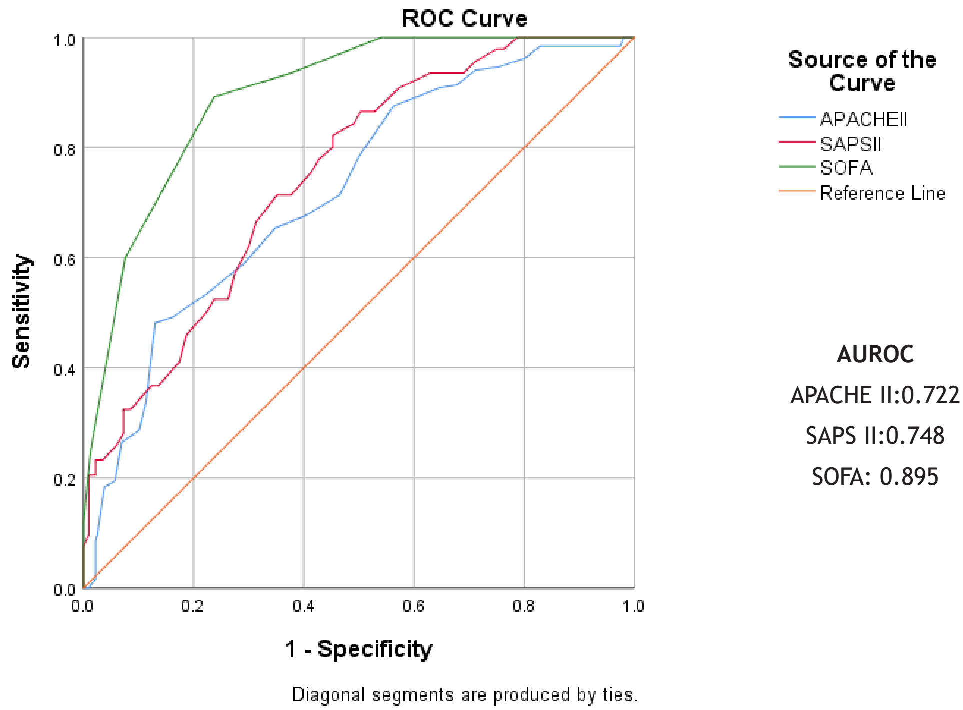


**Figure 1. Comparison of the Area under Receiver operating characteristic curve of the APACHE II, SAPS II and SOFA scoring system for outcome prediction of critically ill patient in intensive care unit.**

**Table 2. The Hosmer-Lemeshow (H-L) test for the goodness-of-fit evaluation for APACHE II, SAPS II and SOFA scoring system.**

| Scoring System | H-L Chi-square | p-value |
|---|---|---|
| APACHE II | 18.757 | 0.016 |
| SAPS II | 22.376 | 0.004 |
| SOFA | 9.907 | 0.129 |

APACHE II: Acute physiology and chronic health evaluation II; SAPS II: Simplified acute physiology score II; SOFA: sequential organ failure assessment score

Better fitted model was further assessed with logistic regression model for predictive performance assessment for the SOFA scoring system. A negative coefficient (B= -1.094) indicates that as the SOFA score increases, the probability of survival decreases. The odds ratio (OR=0.335; 95% CI=0.270-0.415) suggests that for every unit increase in SOFA score, the odds of survival decrease by 66.5% which was statistically significant ($p < 0.001$). Statistically significant result indicating a strong association. The confidence interval does not include 1, confirming that the SOFA score is a strong predictor of mortality (Table 3).

**Table 3. Predictive performance assessment for the SOFA severity scoring system.**

| Scoring system | B | Exp (B) Odds ratio | Significance p | 95% CI for EXP(B) | |
|---|---|---|---|---|---|
| | | | | Lower | Upper |
| SOFA | -1.094 | .335 | <0.001 | .270 | .415 |

SOFA: sequential organ failure assessment score; OR=odds ratio

The SOFA score shows the best model fit among the three scoring systems (p = 0.129). SOFA score is a strong predictor of mortality, with higher scores indicating increased risk of death. APACHE II and SAPS II do not fit the data well (p-values < 0.05), suggesting they may not be as reliable for this population (Table 2).

The Standardized Mortality Ratio (SMR) is a measure of how well a severity scoring system predicts mortality compared to observed outcomes. APACHE II and SAPS II indicating moderate Calibration. Both APACHE II (SMR = 0.87, p = 0.431) and SAPS II (SMR = 0.88, p = 0.517) slightly overestimate mortality (predicting higher death rates than observed). Their p-values (> 0.05) indicate no significant difference between observed and predicted mortality, suggesting reasonable but not perfect calibration. The confidence intervals include 1.0, meaning these models do not significantly misrepresent mortality. SOFA (SMR = 0.38, p < 0.001) heavily overestimates mortality, predicting a much higher death rate (95.2%) than observed (36.9%). The low p-value (< 0.001) suggests a significant difference between observed and predicted mortality, indicating poor calibration. The confidence interval (0.282 - 0.511) does not include 1.0, confirming that SOFA systematically overestimates mortality. APACHE II and SAPS II show moderate calibration, with no significant difference between predicted and actual mortality. SOFA overestimates mortality significantly, suggesting it may not be as reliable for absolute mortality predictions in this cohort (Table 4).

**Table 4. Standardized mortality ratio for the APACHE II, SAPS II and SOFA Scoring system. n=501**

| Scoring system | Observed mortality | Predicted mortality | SMR | p-value | 95% CI | |
|---|---|---|---|---|---|---|
| | | | | | Lower | Upper |
| APACHE II | 36.9% | 42.40% | 0.87 | 0.431 | 0.634 | 1.148 |
| SAPS II | 36.9% | 41.50% | 0.88 | 0.517 | 0.648 | 1.173 |
| SOFA | 36.9% | 95.2% | 0.38 | <0.001 | 0.282 | 0.511 |

APACHE, Acute physiology and chronic health evaluation; SAPS, Simplifies acute physiology score, and SOFA, sequential organ failure assessment score, and SMR: Standardized mortality ratio

The Mann-Whitney U test was used to compare the median scores of APACHE II, SAPS II, and SOFA between survivors and non-survivors, determining if there is a significant difference in severity scores between the two groups. Severity Scores are significantly higher in non-survivors. APACHE II median score is 25 in non-survivors vs. 19.5 in survivors (p = < 0.001). SAPS II median score is 55 in non-survivors vs. 42 in survivors (p= < 0.001). SOFA median score is 15 in non-survivors vs. 12 in survivors (p = < 0.001). Non-survivors consistently have significantly higher severity scores than survivors, indicating that all three scoring systems effectively differentiate between survivors and non-survivors. SOFA has the largest difference between groups. The mean rank for SOFA in non-survivors (375.77) is much higher than in survivors (177.95). This suggests that SOFA may be the strongest differentiator between survivors and non-survivors compared to APACHE II and SAPS II. While all models show significant differences between survivors and non-survivors, SOFA has the most pronounced difference, making it a potentially stronger predictor of mortality. The Mann-Whitney U test values for all three models are highly significant (p < 0.001). This confirms that the differences in severity scores between survivors and non-survivors are not due to chance. All three scoring systems are statistically significant in differentiating between survival and mortality outcomes in critically ill patients (Table 5).

**Table 5. Comparison of median APACHE II, SAPS II and SOFA Score among Survivors and non-survivors n=501**

| Outcome | Number | Median | Mean Rank | Mann-Whitney U test | p-value |
|---|---|---|---|---|---|
| **APACHE II** | | | | | |
| Non-survivor | 185 | 25 | 321.02 | 16275.5 | <0.001 |
| Survivor | 316 | 19.50 | 210.00 | | |
| **SAPS II** | | | | | |
| Non-survivor | 185 | 55 | 329.29 | 14746.0 | <0.001 |
| Survivor | 316 | 42 | 205.16 | | |
| **SOFA** | | | | | |
| Non-survivor | 185 | 15 | 375.77 | 6147.0 | <0.001 |
| Survivor | 316 | 12 | 177.95 | | |

## DISCUSSION

This study evaluated the predictive performance of APACHE II, SAPS II, and SOFA among critically ill patients admitted to adult ICUs of a teaching hospital in Nepal. The observed ICU mortality rate was 36.9%, comparable to reports from Nepal, India, and Poland. [2,14,15] Higher mortality rates, up to 50%, have been documented in some Indian and Vietnamese studies, [16-18] whereas lower rates are commonly reported in developed countries.[19-21] These differences may reflect variations in case mix, disease severity, healthcare infrastructure, timeliness of ICU admission, workforce expertise, and resource availability. Discriminatory ability assessed using ROC analysis demonstrated that SOFA had excellent performance (AUC 0.895), outperforming SAPS II (0.748) and APACHE II (0.722). This finding aligns with several studies confirming strong SOFA discrimination,[21,22] although some investigations have reported lower predictive ability.[18,23] Such inconsistencies may be attributed to heterogeneity in patient populations, clinical settings, and methodological differences.

Calibration assessed by the Hosmer-Lemeshow test showed that SOFA provided the best model fit ($p = 0.129$), while APACHE II ($p = 0.016$) and SAPS II ($p = 0.004$) demonstrated poor fit. Many published studies have not reported Hosmer-Lemeshow results for SOFA,[15,16,18,21-26] and calibration findings for APACHE II [16,18,19,23,25,27] and SAPS II [23,27-29] are inconsistently documented, with only a few studies providing comprehensive calibration analyses.[14,17,20,30,31] These underreporting limits meaningful comparison across settings. Standardized mortality ratio (SMR) analysis further evaluated calibration. APACHE II and SAPS II showed moderate calibration without significant differences between predicted and observed mortality, whereas SOFA significantly overestimated mortality (SMR 0.38, $p < 0.001$). While SOFA demonstrated excellent discrimination and acceptable Hosmer-Lemeshow fit, its SMR findings suggest reduced precision in absolute mortality prediction in this cohort. Previous literature rarely reports SMR for SOFA, and only limited studies have assessed SMR for APACHE II and SAPS II. [2,20,31]

The Mann-Whitney U test revealed significantly higher median scores among non-survivors across all three systems ($p < 0.001$), with SOFA demonstrating the most distinct separation. Few studies have applied this comparative approach, [14] yet it provides additional insight into the practical differentiation of outcomes. Collectively, evaluation across discrimination, calibration, and group comparison parameters indicates that SAPS II performed better than APACHE II, while SOFA emerged as the strongest overall predictor in this ICU setting. Despite their utility, these scoring systems have limitations. They may not account for lead-time bias, pre-ICU management factors, evolving clinical interventions, or contextual variables such as staffing patterns, nurse-to-patient ratios, laboratory support, and financial constraints. Inter-observer variability may also affect scoring accuracy.

## CONCLUSIONS

All three scoring systems effectively differentiated survivors from non-survivors. Among them, SOFA exhibited superior discriminative performance and the best overall predictive accuracy, although it tended to overestimate absolute mortality. SAPS II and APACHE II showed moderate discrimination and calibration but performed less robustly compared with SOFA.

## ACKNOWLEDGEMENTS

We sincerely acknowledge the valuable contributions and support of all study participants, the ICU In-charge, and the dedicated ICU staff.

## CONFLICT OF INTEREST

Authors declare there are no any conflicts of Interest.

## REFERENCES

1. Udwadia FE. Principles of critical care. 3rd ed. Mumbai: Jaypee Digital; 2014. p.7-13. https://www.jaypeedigital.com/book/9789351522355

2. Mehta RK, Qureshi A, Arora M. A study to evaluate predictors of mortality in mechanically ventilated patients using APACHE II and SAPS II. Connect World Crit Care Nurs. 2014;9(3):110-4. [Article]

3. Wong DT, Knaus WA. Predicting outcome in critical care: the current status of the APACHE prognostic scoring system. Can J Anaesth. 1991;*38*(3):374-83. [Article]

4. Hunt JP, Meyer AA. Predicting survival in the intensive care unit. Curr Probl Surg. 1997;34(7):527-99. [Article]

5. 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(8):591-7.[PubMed]

6. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II. Crit Care Med. 1985;13(10):818-29. [Article]

7. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Chest. 1991;100(6):1619-36. [PubMed]

8. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute physiology and chronic health evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. Crit Care Med. 2006;34(5):1297-310.[PubMed]

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(11):975-7. [PubMed]

10. Le Gall JR. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. JAMA. 1993;270(24):2957-63. [Article]

11. Vazquez G, Benito S, Rivera R. Simplified acute physiology score III: a project for a new multidimensional tool for evaluating intensive care unit performance. Crit Care. 2003;7(5):345-6. [Article]

12. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. Intensive Care Med. 1996;22:707-10.[PubMed]

13. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inform. 2014;48:193-204.[PubMed]

14. Parajuli BD, Shrestha GS, Pradhan B, Amatya R. Comparison of acute physiology and chronic health evaluation II and acute physiology and chronic health evaluation IV to predict intensive care unit mortality. Indian J Crit Care Med. 2015;19(2):87-91. [Download PDF]

15. Fuchs PA, Czech IJ, Krzych ŁJ. The pros and cons of the prediction game: the never-ending debate of mortality in the intensive care unit. Int J Environ Res Public Health. 2019;16(18):3394.[Article]

16. Asmarawati TP, Suryantoro SD, Rosyid AN, Marfiani E, Windradi C, Mahdi BA, et al. Predictive value of sequential organ failure assessment, quick sequential organ failure assessment, acute physiology and chronic health evaluation II, and new early warning signs scores estimate mortality of COVID-19 patients requiring intensive care unit. Indian J Crit Care Med. 2022;26(4):464-71.[Article]

17. Nagar VS, Sajjan B, Chatterjee R, Parab NM. The comparison of APACHE II and APACHE IV score to predict mortality in intensive care unit in a tertiary care hospital. Int J Res Med Sci. 2019;7(5):1598-603.[Article]

18. Do SN, Dao CX, Nguyen TA, Nguyen MH, Pham DT, Nguyen NT, et al. Sequential organ failure assessment (SOFA) score for predicting mortality in patients with sepsis in Vietnamese intensive care units: a multicentre, cross-sectional study. BMJ Open. 2023;13(3):1-14.

19. Sadaka F, Ethmane Abou Eimaali C, Cytron MA, Fowler K, Javaux VM, O'Brien J. Predicting mortality of patients with sepsis: a comparison of APACHE II and APACHE III scoring systems. J Clin Med Res. 2017;9(11):907-10.[Article]

20. Katsounas A, Kamacharova I, Tyczynski B, Eggebrecht H, Erbel R, Canbay A, et al. The predictive performance of the SAPS II and SAPS 3 scoring systems: a retrospective analysis. J Crit Care. 2016;33:180-5.[PubMed]

21. Fayed M, Patel N, Angappan S, Nowak K, Vasconcelos Torres F, Penning DH, et al. Sequential organ failure assessment (SOFA) score and mortality prediction in patients with severe respiratory distress secondary to COVID-19. Cureus. 2022;14(7):1-8. [PubMed]

22. Minne L, Abu-Hanna A, de Jonge E. Evaluation of SOFA-based models for predicting mortality in the ICU: a systematic review. Crit Care. 2008;12(6):1-13.[Article]

23. Schoe A, Bakhshi-Raiez F, De Keizer N, Van Dissel JT, De Jonge E. Mortality prediction by SOFA score in ICU patients after cardiac surgery: comparison with traditional prognostic models. BMC Anesthesiol. 2020;20(1):1-8.[PubMed]

24. Pölkki A, Pekkarinen PT, Takala J, Selander T, Reinikainen M. Association of sequential organ failure assessment (SOFA) components with mortality. Acta Anaesthesiol Scand. 2022;66(6):731-41.[PubMed]

25. Bajpai J, Prakash V, Kant S, Verma AK, Srivastava A, Bajaj DK, et al. Study of oxidative stress biomarkers in chronic obstructive pulmonary disease and their correlation with disease severity in North Indian population cohort. Lung India. 2017;34(4):324-9. [PubMed]

26. Saleh A, Ahmed M, Sultan I, Abdel-lateif A. Comparison of the mortality prediction of different ICU scoring systems (APACHE II and III, SAPS II, and SOFA) in a single center ICU subpopulation with acute respiratory distress syndrome. Egypt J Chest Dis Tuberc. 2015;64(4):843-8.[Article]

27. Paudel P, Rai S, Shrestha S, Pradhan GBN, Bhattachan CL. Analysis of outcomes of critically ill surgical patients using SAPS II score. J Inst Med Nepal. 2019;41(1):85-90. [Article]

28. Singh P, Pathak S, Sharma R. A comparison of acute physiology and chronic health evaluation III and simplified acute physiology score II in predicting sepsis outcome in intensive care unit. Anesth Essays Res. 2018;12(2):592-7.[PubMed]

29. Moon BH, Park SK, Jang DK, Jang KS, Kim JT, Han YM. Use of APACHE II and SAPS II to predict mortality for hemorrhagic and ischemic stroke patients. J Clin Neurosci. 2015;22(1):111-5.[PubMed]

30. Faruq MO, Mahmud MR, Begum T, Ahsan AA, Fatema K, Ahmed F, et al. A comparison of severity systems APACHE II and SAPS II in critically ill patients. Bangladesh Crit Care J. 2013;1(1):27-32.[Article]

31. Poncet A, Perneger TV, Merlani P, Capuzzo M, Combescure C. Determinants of the calibration of SAPS II and SAPS 3 mortality scores in intensive care: a European multicenter study. Crit Care. 2017;21(1):2-10.[PubMed]