Comparison of acute physiology, age, chronic health evaluation III score with initial sequential organ failure assessment score to predict ICU mortality

GS Shrestha, R Gurung and R Amatya

Department of Anaesthesiology, Institute of Medicine, Tribhuvan University Teaching Hospital, Maharajgunj, Kathmandu, Nepal

Corresponding author: Dr. Gentle Sunder Shrestha, MD Anaesthesiology, Lecturer, Department of Anaesthesiology, Tribhuvan University Teaching Hospital, Maharajgunj, Kathmandu, Nepal; e-mail: gentlesunder@hotmail.com

ABSTRACT

Critically ill patients are provided with highest level of monitoring, care and treatment in Intensive Care Unit (ICU), which is very expensive and consumes many hospital resources. Various scoring systems have been developed to predict outcome in ICU patients so as to help physicians to prioritize patient admission and management. The objective of this study was to compare Acute Physiology and Chronic Health Evaluation (APACHE) III score with initial Sequential Organ Failure Assessment (SOFA) score to predict ICU mortality. Hundred seventeen patients admitted consecutively in ICU were enrolled. APACHE III and initial SOFA score of individual patients were calculated based on worst values in first 24 hours of admission. Outcome was recorded as survivors or non survivors in ICU. Both the scores were significantly higher in non survivors ($p<0.001$). A positive and strong correlation was seen between the scores with Spearman's rho correlation coefficient of 0.866 ($p<0.001$). Discrimination for APACHE III and initial SOFA score was good with area under ROC curve of 0.895 and 0.879 respectively. Cut off point with best Youden index was e" 61 for APACHE III and e" 8 for initial SOFA score. ICU mortality differed significantly above and below cut off points ($p<0.001$). Hosmer Lemeshow test showed initial SOFA score to have better calibration than APACHE III score. Initial SOFA score is comparable to APACHE III score for mortality prediction in ICU and so can be helpful for better utilization of limited resources in ICU.

Keywords: APACHE III score, initial SOFA score, ICU mortality.

INTRODUCTION

Critical care is very expensive. Intensive Care Unit (ICU) beds are limited and constitute only 8-10% of all hospital beds, yet accounts for a major part of hospital expenditure.[1] Scoring systems are designed to objectively quantify physiologic derangements and comorbid conditions for estimating mortality, length of stay and ICU resource use.[2] Precise disease classification and accurate outcome prediction can optimize ICU bed usage by reducing unnecessary low-risk monitored-only patients and futile care of terminally ill patients.[3]

APACHE (Acute Physiology, Age, Chronic Health Evaluation) III was introduced in 1991 with a much larger database and better predictive capacity than APACHE II. APACHE III score is the sum of acute physiology score, age score and chronic health problem score. Acute physiology score is based on the worst physiological values during the first 24 hours of admission. Scores range from 0 to 299 (acute physiology 0 to 252, chronic health evaluation 0 to 23 and age 0 to 24) with higher values having the worst prognosis, as described by Knaus *et al*.[4] APACHE III accurately predicted ICU mortality in United States,[5] Australia,[6] Brazil[7] and Germany.[8] Performance of APACHE III was better than other scoring systems in some studies,[9-11] but this system is complex, difficult to administer and proprietary.[2]

Sequential Organ Failure Assessment (SOFA) scoring system was introduced by European Society of Critical Care Medicine (ESCCM) in 1994. It evaluates six different organ systems based on simple and routinely available variables (PaO_2/FiO_2 for respiratory system, mean arterial pressure for cardiovascular system, Glasgow coma scale for central nervous system, serum creatinine or urine output for renal system, platelet count for coagulation system and serum bilirubin for hepatic

Table-1: Distribution of age, duration of stay, APACHE III sore and initial SOFA score.

	Mean	Standard Deviation	Minimum	Maximum
Age (years)	43.18	18.49	16	84
Duration of stay (days)	11.33	11.25	1	66
APACHE III score	66.99	33.49	16	162
Initial SOFA score	8.38	4.28	0	23

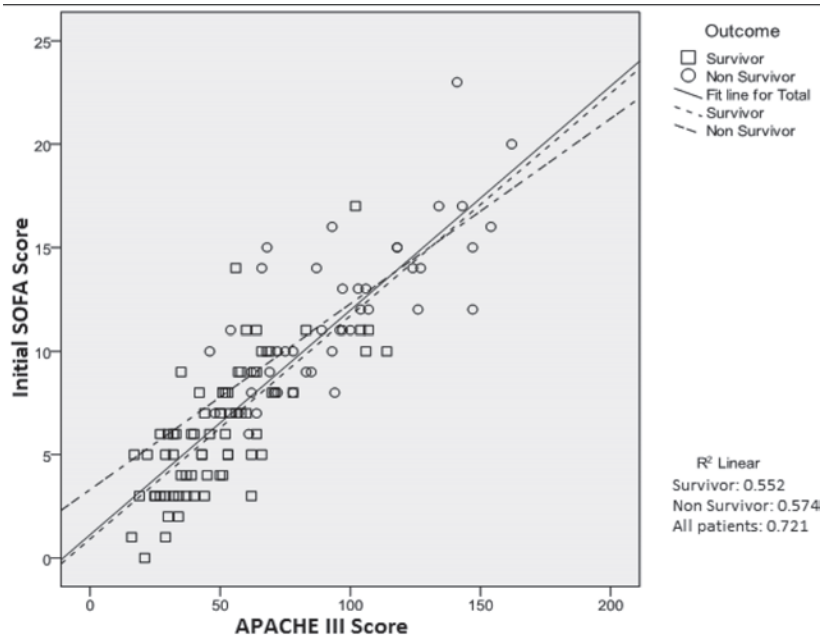

Fig. 1. Correlation between APACHE III and initial SOFA score. Spearman's rho correlation coefficient (r^2) is 0.721 for all patients, 0.552 for survivors and 0.574 for non survivors.

system). All parameters are graded 0-4, which are summed up to calculate SOFA score. It also incorporates therapeutic interventions like mechanical ventilation and use of inotropes. Score can be calculated daily in ICU which takes into consideration the changing severity of organ dysfunction over time as described by Vincent *et al*.[12] Initial SOFA score is calculated based on worst values in 24 hours immediately following admission.[13-15] Some studies have found initial SOFA score to be a good predictor of outcome in ICU.[9,14,16,17]

Our study aims to see whether initial SOFA score can predict ICU mortality as effectively as APACHE III score.

MATERIALS AND METHODS

Data was prospectively collected from six bed multidisciplinary ICU. A total of 117 consecutively admitted ICU patients from September 2009 to March 2010 were enrolled. Patients were excluded if age was less than 16 years or the patients were taken away from ICU against medical advice. Laboratory reports and clinical information necessary for APACHE III and SOFA scoring were obtained and the scores were calculated based on worst values in first 24 hours of ICU admission. APACHE III and initial SOFA score were calculated as defined in original reports.[4,13] Outcome was recorded as survivors or non survivors in ICU.

Statistical Analysis: Data were analyzed using descriptive statistics, frequency distribution, independent t test, chi square test, scatter diagram, linear regression analysis and univariate binary logistic regression analysis. Variables found significant in univariate analysis were analyzed using multivariate analysis. Spearman's rho test was used to calculate correlation between APACHE III and initial SOFA score. Discrimination was tested using the area under receiver-operating characteristic (ROC) curve. ROC analysis was also performed to calculate the cut-off values, sensitivity, specificity and overall correctness of prediction. The best Youden index (sensitivity + specificity – 1) was used to determine the best cut-off point. Survivors and non-survivors were compared above and below the cutoff points. Calibration, which compares the number of observed and predicted deaths, was assessed using the Hosmer-Lemeshow goodness-of-fit test. Data were entered in Microsoft Excel 2003 and analyzed using SPSS program, version 17.0.

Sample size calculation: Sample size (117 patients) was calculated to ensure power of 0.80 using the formula z^2pq/d^2. Pretest of 60 cases showed proportion of non survivors (p) to be 0.45, proportion of survivors (q) to be 0.55 with maximum tolerable error (d) of 0.09 and reliability coefficient for 95% confidence interval of 1.96.

RESULTS

Mean age was 43.18±18.49 years and 73 (62.4%) of the

Table-2: Comparison of age, duration of stay, APACHE III and initial SOFA score for survivors and non survivors.

	Outcome	Mean	Standard Deviation	Standard Error of Mean	p Value
Age	Survivor	41.11	17.332	2.029	0.11
	Non Survivor	46.61	19.998	3.015	
Duration of stay	Survivor	11.10	10.516	1.231	0.77
	Non Survivor	11.73	12.483	1.882	
APACHE III score	Survivor	50.14	21.762	2.547	<0.001
	Non Survivor	94.95	30.836	4.649	
Initial SOFA score	Survivor	6.32	3.157	0.370	<0.001
	Non Survivor	11.82	3.649	0.550	

Table-3: Prediction of ICU mortality on the first day of ICU admission.

	Cut off point	Sensitivity (%)	Specificity (%)	Youden index	Overall correctness (%)
APACHE III score	≥61	90.91	73.97	0.65	80.34
Initial SOFA score	≥8	90.91	65.75	0.57	75.21

patients were male. ICU mortality was 37.6%. Demographic data, APACHE III score and initial SOFA score are listed in Table-1. APACHE III score and initial SOFA score were compared between survivors and non survivors as shown in Table-2 and both the scores were significantly higher in non survivors (p<0.001). Age and duration of ICU stay did not differ significantly between survivors and non survivors (Table-2). For initial SOFA score, with unit increase in score, there was 1.645 (95% CI=1.367, 1.979) times higher odds for mortality and for APACHE III score, there was 1.066 (95% CI=1.042, 1.091) times higher odds for mortality. Univariate analysis revealed that females were more likely to be non survivors (p=0.01). Patients requiring mechanical ventilation in the first 24 hours of ICU admission were more likely to die (p<0.001). Similarly, patients requiring inotropic support during first 24 hours had the higher chance of being the non survivors (p=0.01). However, multivariate analysis showed only APACHE III score (p=0.031; OR=1.036, 95% CI=1.003, 1.070) and initial SOFA score (p=0.024; OR=1.359, 95% CI=1.041, 1.773) to have statistically significant relationship with outcome.

A positive and strong correlation was seen between initial SOFA score and APACHE III score. Spearman's rho correlation coefficient (r^2) was 0.721 for all patients indicating that 72.1% variance in initial SOFA score is

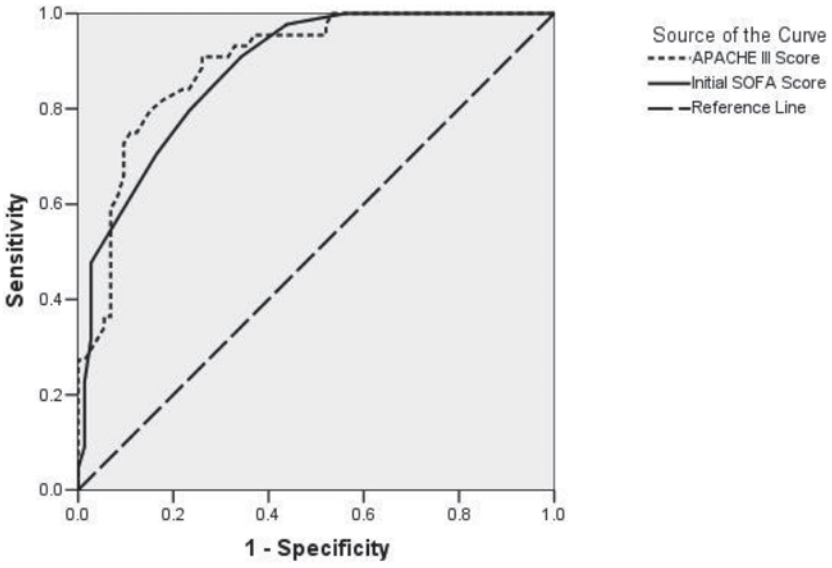

Fig. 2. ROC curves for APACHE III and initial SOFA score. The area under curve is 0.895 for APACHE III and 0.881 for initial SOFA score.

Table-4: Hosmer-Lemeshow goodness-of-fit statistics for APACHE III and initial SOFA score.

Predicted deciles of mortality (%)	APACHE III score					Initial SOFA score				
	n	Survivor		Non survivor		*n*	Survivor		Non survivor	
		Observed	Expected	Observed	Expected		Observed	Expected	Observed	Expected
0-10	12	12	11.594	0	0.406	16	16	15.579	0	0.421
>10-20	12	12	11.290	0	0.710	16	16	14.873	0	1.127
>20-30	12	11	10.780	1	1.220	10	9	8.722	1	1.278
>30-40	12	11	10.043	1	1.957	10	7	8.057	3	1.943
>40-50	11	8	8.522	3	2.478	13	8	9.308	5	3.692
>50-60	12	8	8.355	4	3.645	9	5	5.446	4	3.554
>60-70	13	5	7.419	8	5.581	10	5	4.823	5	5.177
>70-80	12	1	3.432	11	8.568	14	5	4.640	9	9.360
>80-90	12	5	1.397	7	10.603	12	1	1.371	11	10.629
>90-100	9	0	0.167	9	8.833	7	1	0.181	6	6.819
$\Sigma(E\text{-}O)^2/E$	χ^2=16.904, df=8, p=0.031					χ^2=7.140, df=8, p=0.522				

n, number of patients per decile; *E*, expected number of deaths; *O*, observed number of deaths; df, degrees of freedom.

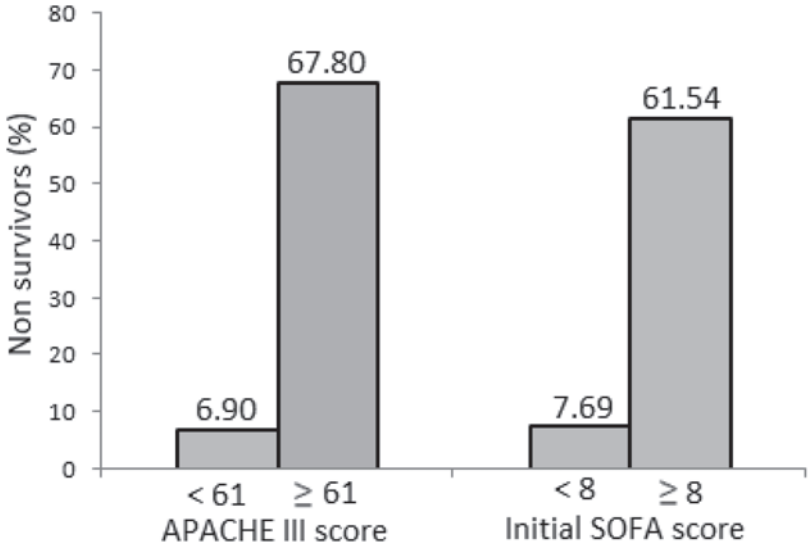

Fig. 3. Non survivors above and below the cut off point giving the best Youden index for APACHE III and initial SOFA score

explained by APACHE III score. The coefficient was 0.552 for survivors and 0.574 for non survivors (Fig. 1).

Area under ROC curve for APACHE III score was 0.895 (p<0.001; 95% CI=0.839, 0.952) and for initial SOFA score it was 0.881 (p<0.001; 95% CI=0.822, 0.940) as shown in Fig. 2. Discrimination was good for both APACHE III and initial SOFA score.

Calculation of Youden index showed the best cut off point for APACHE III to be e" 61 and the best point to be e" 8 for initial SOFA score. At these cut off points, APACHE III had Youden index of 0.65 and 80.3% overall correctness of prediction and initial SOFA score had Youden index of 0.57 with overall correctness of prediction of 75.2% (Table-3). Among the patients with APACHE III score < 61, 6.9% were non survivors whereas 67.8% of patients with APACHE III score e" 61 died (p<0.001). Similarly, as shown in Fig. 3, among the patients with initial SOFA score < 8, 7.7% died whereas 61.5% of patients with initial SOFA score e" 8 were non survivors (p<0.001).

For assessing goodness-of-fit, Hosmer and Lemeshow test for APACHE III score produced ÷ ² of 16.904 and p of 0.031 indicating the model does not fit the data. As shown in Table-4, initial SOFA score produced ÷ ² of 7.140 and p of 0.522. For initial SOFA score, which produced insignificant ÷ ², the model is good and has better calibration than APACHE III score. Initial SOFA score correctly predicted 70.5% non survivors and 83.6% survivors. Similarly, APACHE III score correctly predicted 65.9% non survivors and 90.4% survivors.

DISCUSSION

APACHE III and SOFA score were shown to perform well in a variety of patient populations.[9,18-20] Medline search did not show any study in multidisciplinary ICU comparing these two scoring systems. So we compared initial SOFA score (simple, with few variables and economical) with APACHE III score (based on large database with enhanced predictive capacity) in our multidisciplinary ICU enrolling both medical and surgical patients.

In our study, there was no significant relationship between age of patient and outcome (p=0.11) as shown in Table-2. It is consistent with the findings of Acharya *et al*.[17] Influence of age on outcome was shown to decrease with increasing disease severity.[21] Similarly, there was no relation between duration of ICU stay and outcome (p=0.77). It was in contrary to studies by Acharya *et al*[17] and Schuster *et al*[22] where non survivors had shorter duration of ICU stay. This might be because our study also enrolled postoperative patients admitted in ICU for short period of observation.

Both mean APACHE III and initial SOFA score were significantly (p<0.001) higher in non survivors when compared to survivors (Table 2). Similar results were seen in studies by Ferreira *et al*[14], Acharya *et al*[17] and Chen *et al*.[9] os A positive and strong correlation was seen between initial SOFA score and APACHE III score (r^2 of 0.721 for all patients) (Fig. 1). Similar correlation was observed in a study by Chen *et al*[9] (r^2 of 0.628 for all patients). Discrimination was good for both APACHE III (area under ROC curve 0.895) and initial SOFA score (area under ROC curve 0.881) (Fig. 2). Similar results were seen in other studies. Area under ROC curve for APACHE III was 0.90 in a study by Knaus *et al*[4] and 0.89 in a study by Zimmerman *et al*.[5] Area for initial SOFA was 0.917 in a study by Chen *et al*[9] and 0.79 in a study by Ferreira *et al*.[14] There is a significant difference (p<0.001) in non survivors above and below the best cut off point giving the highest Youden index for both APACHE III and initial SOFA score (Fig. 3). Hosmer Lemeshow test showed initial SOFA score to produce insignificant ÷ ² value and thus had a better calibration and performed better to predict non survivors when compared with APACHE III score (Table 4). Similar results were seen in a study by Chen *et al*[9] where initial SOFA score (÷ ² = 5.006, eight degrees of freedom [df], p = 0.757) had better calibration than APACHE III score (÷ ² = 10.392, eight degrees of freedom [df], p = 0.239) in cirrhotic patients.

Despite the encouraging results, our study has some limitations. First, this study was conducted in a single

center and enrolled both surgical and medical patients. So the results may not be generalized to other centers or to the ICUs dedicated specially for management of medical or surgical patients. Second, only initial SOFA score was calculated. Daily SOFA scoring would further enhance the predictive capacity. Finally, patients were followed only till ICU discharge. Larger multicentered studies and evaluation of special category of patients may be helpful.

In conclusion, this study demonstrates that there is a strong and good correlation between APACHE III and initial SOFA score. Discrimination was good for both the scores. Moreover, initial SOFA score had better calibration and performed better to predict non survivors when compared with APACHE III score. So initial SOFA score can be used as a simple, economical yet reliable tool to predict outcome in ICU and can help clinicians for better utilization of limited and expensive ICU resources.

REFERENCES

1. Critical care. In Morgan GE, Mikhail MS, Murray MJ, editors. Clinical Anesthesiology (4th ed.). United States of America: McGraw-Hill Companies 2006: 1018-64.
2. Liu L, Grooper MA. Overview of Anesthesiology and Critical Care Medicine. In Miller RD, editor. Miller's Anesthesia (7th ed.). United States of America: Churchill Livingstone 2010: 2845-67.
3. Wong DT, Knaus WA. Predicting outcome in critical care: the current status of the APACHE prognostic scoring system. *Canadian J Anaesth* 1991; 38: 374-83.
4. Knaus WA, Wagner DP, Draper EA *et al*. The APACHE III Prognostic System. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991; 100: 1619-36.
5. Zimmerman JE, Wagner DP, Draper EA, Wright LRN, Alzola CMS, Knaus WA. Evaluation of Acute Physiology and Chronic Health Evaluation III prediction of hospital mortality in an independent database. *Crit Care Med* 1998; 26: 1317-26.
6. Cook DA. Performance of APACHE III models in an Australian ICU. *Chest* 2000; 118: 1732-8.
7. Bastos PG, Sun X, Wagner DP, Knaus WA, Zimmerman JE. Application of the APACHE III prognostic system in Brazilian intensive care units: A prospective multicenter study. *Intensive Care Med* 1996; 22: 564-70.
8. Von Bierbrauer A, Riedel S, Cassel W, Von Wichert P. Validierung des Acute Physiology and Chronic Health Evaluation (APACHE) III Scoring systems. *Der Anaesthesist* 1998; 47: 30-38.
9. Chen YC, Tian YC, Liu NJ *et al*. Prospective cohort study comparing sequential organ failure assessment and acute physiology, age, chronic health evaluation III scoring systems for hospital mortality prediction in critically ill cirrhotic patients. *Int'l J Clin Pract* 2006; 60: 160-66.
10. Butt AK, Khan AA, Alam A, Shah SWH, Shafqat F, Naqvi AB. Predicting hospital mortality in cirrhotic patients: comparison of child-pugh and acute physiology, age and chronic health evaluation (APACHE III) scoring systems. *Am J Gastroenterol* 1998; 93: 2469-75.
11. Reina A, Vazquez G, Aguayo E, Bravo I, Colmenero M, Bravo M. Mortality discrimination in acute myocardial infarction: comparison between APACHE III and SAPS II prognosis systems. *Intensive Care Med* 1997; 23: 326-30.
12. Vincent JL, Moreno R, Takala J *et al*. The SOFA (Sepsis related organ failure assessment) score to describe organ dysfunction or failure. On behalf of the working group on sepsis related problem of the European society of Intensive care medicine. *Intensive Care Med* 1996; 22: 707-710.
13. Vincent JL, de Mendonca A, Cantraine F *et al*. Use of the SOFA score to assess the incidence of organ dysfunction/ failure in intensive care units: Results of a multicenter, prospective study. *Crit Care Med* 1998; 26: 1793-800.
14. Ferreira FL, Bota DP, Bross A, Melot C, Vincent JL. Serial evaluation of the SOFA Score to predict outcome in critically ill patients. *J Amer Med Assoc* 2001; 286: 1754-8.
15. Ceriani R, Mazzoni M, Bortone F *et al*. Application of the Sequential Organ Failure Assessment Score to Cardiac Surgical Patients. *Amer Coll Chest Phys* 2003; 123: 1229-39.
16. Kim SH, Lee MG, Park SM *et al*. The SOFA Score to Evaluate Organ Failure and Prognosis in the Intensive Care Unit Patients. *Tuberc Respir Dis* 2004; 57: 329-35.
17. Acharya SP, Pradhan B, Marhatta MN. Application of "the Sequential Organ Failure Assessment (SOFA) score" in predicting outcome in ICU patients with SIRS. *Kathmandu Univ Med J* 2007; 5: 475-83.
18. Cooke CR, Kahn JM, Caldwell E *et al*. Predictors of hospital mortality in a population-based cohort of patients with acute lung injury. *Crit Care Med* 2008; 36: 1412-20.
19. Krenn H, Dorn C, Fitzgerald RD, Pojer H, Seitlinger GR, List WF. Prognostic value of APACHE II and III for preoperative evaluation of emergency surgery patients. *Eur J Anaesthesiol* 2000; 17: 529-30.
20. Ahlstrom A, Kuitunen A, Peltonen S *et al*. Comparison of 2 acute renal failure severity scores to general scoring systems in the critically ill. *Amer J Kidney Dis* 2006; 48: 262-8.
21. Nicolas F, Le Gall JR, Alperovitch A *et al*. Influence of patient's age on survival, level of therapy and length of stay in intensive care unit. *Intensive Care Med* 1987; 13: 9-13.
22. Schuster HP, Wilts S, Ritschel P, Schuster FP. Predictive value of score parameters of the Simplified Acute Physiology Score (SAPS) II for the duration of treatment of intensive care patients. *Wien Klin Wochenschr* 1996; 108: 462-6.