Research Article

Comparison of acute physiology and chronic health evaluation II and acute physiology and chronic health evaluation IV to predict intensive care unit mortality

Bashu Dev Parajuli, Gentle S. Shrestha, Bishwas Pradhan, Roshana Amatya

Abstract

Context: Clinical assessment of severity of illness is an essential component of medical practice to predict the outcome of critically ill-patient. Acute Physiology and Chronic Health Evaluation (APACHE) model is one of the widely used scoring systems. **Aims:** This study was designed to evaluate the Performance of APACHE II and IV scoring systems in our Intensive Care Unit (ICU). **Settings and Design:** A prospective study in 6 bedded ICU, including 76 patients all above 15 years. **Subjects and Methods:** APACHE II and APACHE IV scores were calculated based on the worst values in the first 24 h of admission. All enrolled patients were followed, and outcome was recorded as survivors or nonsurvivors. **Statistical Analysis Used:** SPSS version 17. **Results:** The mean APACHE score was significantly higher among nonsurvivors than survivors ($P < 0.005$). Discrimination for APACHE II and APACHE IV was fair with area under receiver operating characteristic curve of 0.73 and 0.79 respectively. The cut-off point with best Youden index for APACHE II was 17 and for APACHE IV was 85. Above cut-off point, mortality was higher for both models ($P < 0.005$). Hosmer–Lemeshow Chi-square coefficient test showed better calibration for APACHE II than APACHE IV. A positive correlation was seen between the models with Spearman's correlation coefficient of 0.748 ($P < 0.01$). **Conclusions:** Discrimination was better for APACHE IV than APACHE II model however Calibration was better for APACHE II than APACHE IV model in our study. There was good correlation between the two models observed in our study.

Keywords: Acute physiology and chronic health evaluation II, acute physiology and chronic health evaluation IV, intensive care unit mortality

Access this article online
Website: www.ijccm.org
DOI: 10.4103/0972-5229.151016
Quick Response Code:

Introduction

Clinical assessment of severity of illness is an essential component of medical practice, including Intensive Care Unit (ICU) to predict mortality and morbidity of critically ill patient.[1,2] Scoring system can define critically ill-patients, estimate prognosis, guide to allocate the resource and estimate the quality of ICU.[1-3]

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

Correspondence:
Dr. Bashu Dev Parajuli, Department of Anaesthesiology, Institute of Medicine, Tribhuvan University Teaching Hospital, Maharajgunj, Kathmandu, Nepal.
E-mail: basu_parajuli@hotmail.com

Physiological based scoring systems are more applicable than diagnosis based scoring systems and estimate the risk based on the degree of variation from the normal function of major organ systems.[1,2]

Acute Physiology And Chronic Health Evaluation (APACHE), introduced in 1981, takes into consideration of various parameters like physiological variables, vital signs, urine output, neurological score, age and co-morbid conditions, which may have a significant impact on the outcome of Critically ill patients.[4] APACHE II, formulated in 1985, estimate risk based on data available within the first 24 h of admission.[5] APACHE II is a widely used scoring system to quantify the severity of illness in ICU and has been validated in many clinical trials.[5] APACHE IV, introduced in 2006, is the most recent version of APACHE. The new variables added to this model are mechanical ventilation, thrombolysis, impact of sedation on Glasgow coma Scale, Glasgow coma Scale, PaO_2/FiO_2 ratio, pre ICU hospital length of stay, location prior to ICU and 116 disease specific subgroups.[6]

According to the Medline search, no study comparing these 2 scoring systems had been reported in our region (South East Asia). Thus, we designed a study to observe the performance of APACHE II and APACHE IV scoring system in our ICU.

Subjects and Methods

After obtaining consent from institution review board, the study was prospectively conducted for 4 months duration in a 6 bedded ICU of a tertiary level teaching hospital. 76 patients (estimated by using $n = z^2pq/\alpha^2$, where n is sample size, z is the confidence interval (1.96), P is the estimated proportion of attribute present in population (0.3), q = 1 − p and α is tolerable error) all above 15 years, irrespective of diagnosis managed in ICU for > 24 h were enrolled. APACHE II and APACHE IV scores were calculated based on the worst values of the first 24 h of admission. All enrolled patients were followed during their ICU stay and outcome was recorded as survivors or nonsurvivors.

Statistical analysis was carried out using SPSS Version 17 (Polar Engineering and Consulting) and $P < 0.05$ were considered as significant. Descriptive statistics were calculated where applicable. Discrimination is how well a model can predict outcome, was tested by calculating area under receiver operating characteristics (ROC) curve, a graphical plot of true positive (sensitivity) against false positive rate (1-specificity). The best cut-off value was derived by the best Youden Index. Calibration is how well the model tracks the outcome, was tested by Hosmer–Lemeshow Goodness of fit test. Student paired *t*-test was used to compare between the scores. Correlation between the models was calculated by Spearman's rho coefficient.

Results

Among 76 patients, there were 46 male patients and 30 female patients. APACHE II score of the patients ranged from 6 to 35 with a mean of 18.26 ± 7.40. APACHE IV score of the patients ranged from 52 to 151 with a mean of 91.68 ± 22.05. There were 51 survivors (67.1%). For APACHE II model, the mean score for survivors was 16.39 ± 6.82, which was less compared to mean the score of 22.08 ± 7.18 for nonsurvivors (P = 0.001) [Table 1]. For APACHE IV model, the mean score for the survivors was 83.96 ± 17.93, which was less compared with mean the score of 107.44 ± 21.53 for nonsurvivors (P < 0.001) [Table 1]. Mortality increased with increasing APACHE II and APACHE IV score. There was a good correlation between APACHE II and APACHE IV score with Spearman's coefficient of 0.748 (P < 0.01) [Figure 1]. Discrimination for APACHE II and APACHE IV models were fair with area under ROC curve of 0.73 and 0.79 respectively [Figure 2]. The cut-off point with best Youden index for APACHE II was 17 and for APACHE IV was 85. Among patients with APACHE II score ≥ 17, there were 80% nonsurvivors and among patients with APACHE II score < 17, there were 20% nonsurvivors (P = 0.001) [Figure 3]. Similarly, among patients with APACHE IV score ≥ 85, there were 92% nonsurvivors and among patients with APACHE IV score < 85, there were 8% nonsurvivors (P < 0.001) [Figure 4]. There was no linear relationship between predicted length of ICU stay by APACHE IV model and observed ICU length of stay. The Hosmer–Lemeshow Chi-square coefficient value calculated for calibration of APACHE II model was 7.9 (P = 0.34) and of APACHE IV model was 14.26 (P = 0.05).

Table 1: Comparison of mean APACHE II and APACHE IV score among survivors and nonsurvivors

Outcome	Number	Mean score	SD	*P*
APACHE II				
Non survivor	25	22.08	7.18	0.001
Survivor	51	16.39	6.82	
APACHE IV				
Non survivor	25	107.44	21.53	<0.001
Survivor	51	83.96	17.93	

APACHE: Acute Physiology and Chronic Health Evaluation; SD: Standard deviation

Discussion

Both mean APACHE II and APACHE IV score were significantly higher among nonsurvivors than survivors, which correlated with other studies.[5-16]

Acute Physiology and Chronic Health Evaluation II score observed in our study ranged from 6 to 35 with mean score of 18.26 ± 7.40 which was comparable to that reported from India,[7] Bangladesh[8] and Turkey.[9] Survivors had lower mean APACHE II score compared with nonsurvivors, which was statistically significant (P = 0.001) As observed in other studies. Mortality increased with increasing APACHE II score, which was also statistically significant (P = 0.001). Similar results were found in other studies.[7-9]

Acute Physiology and Chronic Health Evaluation IV score ranged from 52 to 151 in our study with a mean of 91.68 ± 22.05, which were comparable to that reported

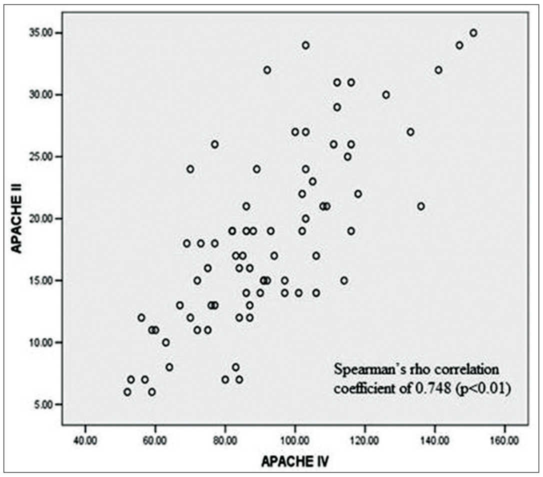

Figure 1: Correlation between Acute Physiology and Chronic Health Evaluation (APACHE) II and APACHE IV score

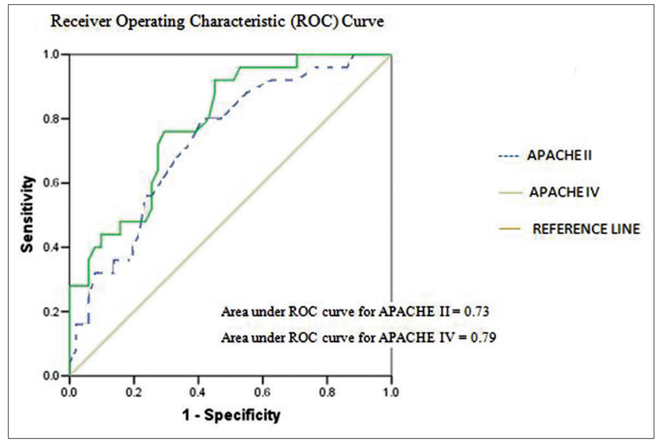

Figure 2: Receiver operating characteristic curve of Acute Physiology and Chronic Health Evaluation (APACHE) II and APACHE IV score to test the discrimination of the model

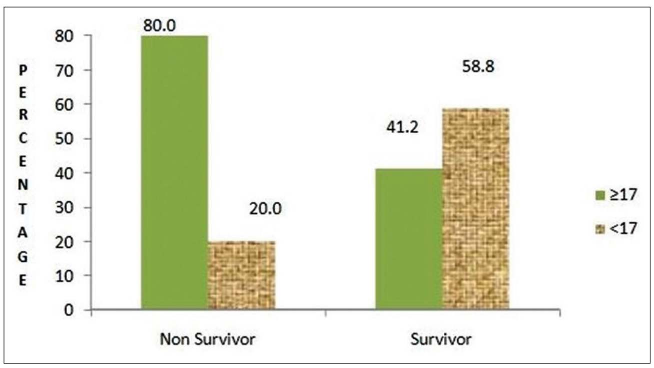

Figure 3: Survivors and nonsurvivors above and below the best cut-off point for Acute Physiology and Chronic Health Evaluation II score

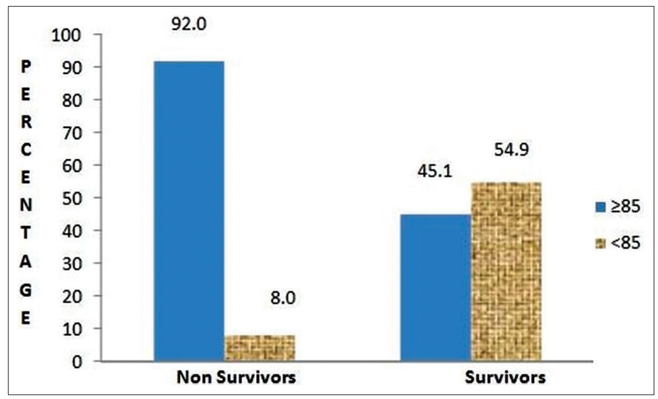

Figure 4: Survivors and nonsurvivors above and below the best cut-off point for Acute Physiology and Chronic Health Evaluation IV score

from Turkey,[9] Karachi.[10] Survivors had lower mean APACHE IV score compared to nonsurvivors, which was statistically significant ($P < 0.001$) as found in other studies. Mortality increased with increasing APACHE IV score, which was also statistically significant ($P < 0.001$) as observed in other studies.[9-16]

The bivariate analysis showed no statistically significant correlation of mortality with age group and sex as observed in other studies. However the association of initial GCS value, mean Billirubin value, Inotropic and mechanical ventilator support [Table 2], high APACHE II score and APACHE IV score were statistically significant with mortality ($P < 0.005$).

There was good correlation between APACHE II and APACHE IV with Spearman's rho correlation coefficient of 0.748 ($P < 0.01$). Similarly correlation among the survivors was 0.708 ($P < 0.01$) and among the nonsurvivors was 0.655 ($P < 0.01$) were also good.

Area under ROC curve observed for APACHE II model was 0.73 which were similar as reported in India,[7] Bangladesh[8] and Turkey.[9] Area under ROC curve observed in our study for APACHE IV was 0.79 which was 0.93 in Ayazoglu study,[9] 0.861 in Keegan *et al*. study[11] and 0.884 in Kramer *et al*. study.[12] The discrimination of APACHE IV model was better than APCHE II model in our study and the finding was consistent with Brinkman *et al*. study[13] but wasn't consistent with Ayazoglu study,[9] Kamal *et al*. study[14] and Lee *et al*. study.[15]

Table 2: Comparison of inotropic and mechanical ventilation support among survivors and nonsurvivors

	Survivors		Non survivors		P
	Number	Percentage	Number	Percentage	
Inotropic support					
Yes	14	48.3	15	51.7	0.006
No	37	78.7	10	21.3	
Mechanical ventilator support					
Yes	28	85.4	21	84	0.001
No	23	45.1	4	16	

Acute Physiology and Chronic Health Evaluation IV model better predict mortality rate than APACHE II scoring system in our ICU. The reason could probably be the consideration of mechanical ventilation support, patient source prior to ICU admission, disease specific subgroup analysis and specific reason for ICU admission.

The Hosmer–Lemeshow Chi-square coefficient for APACHE II was 7.9 (P = 0.34) in our study, which was similar in a study done in Bangladesh[8] and for APACHE IV was 14.26 (P = 0.05) in our study, which was 12.86 (P = 0.12) in a study done in Brazil.[16] Thus, in our study calibration of both models was good however APACHE II had better calibration than APACHE IV model. However, the calibration observed for both models was poor in Brinkman *et al*. study[13] and Lee *et al*. study.[15]

There was uniformity in data collection and investigation to minimize the bias.

Lead time bias is another factor affecting the accuracy of risk prediction,[17] which was difficult to quantify in our study. However, APACHE IV model had considered the location of patient and duration of illness prior to being admitted in ICU.

Besides this, overall quality of ICU is affected by the bed occupancy ratio, lab facility and availability, trained man powers, nurse to patient ratio and financial status of the patient's care giver. Limited resources signifies resources limited within the hospital or outside in the community.[18]

Both APACHE models have shown good performance in the USA[5,6] and other parts of the world. However in the developing countries like ours, the performance of both models can be improved by considering separate points for lead time bias, financial status of the patient's caretakers, nurse to patient ratio in ICU, availability of manpower and resources in the ICU.

Thus, more studies in multiple centers involving larger patient population are needed to validate both the scoring systems in developing countries like ours and separate scoring systems that factor in the pitfalls in resource limited environment need to be developed for good predictability.

To conclude, mortality of patients was significantly high when APACHE II score ≥ 17 and APACHE IV score ≥ 85. Discrimination, was fair for both models, but APACHE IV was superior to APACHE II. Calibration, was better for APACHE II than APACHE IV in our ICU. There was good correlation observed between the models. Further studies with a larger patient population would be required to validate these findings.

Acknowledgment

I would like to acknowledge the support of all ICU staffs, my colleagues, Prof. Dr. Moda Nath Marhatta (Head of Department, Department of Anaesthesiology, TUTH, IOM), Dr. Diptesh Aryal, Dr. Subhash Prasad Acharya (ICU coordinator) and Dr. Amita Pradhan (statistician) to complete this study.

References

1. Palazo M. Severity of illness. In: Vincent JL, Abraham E, Moore FA, Kochanek PM, Fink MP, editors. Text Book of Critical Care. 6th ed. Albany (NY): Elsevier Publishers; 1996. p. 17-29.
2. Bouch DC, Thompson JP. Severity score system in critically ill. Oxf J 2008;8:181-5.
3. Becker RB, Zimmerman JE. ICU scoring systems allow prediction of patient outcomes and comparison of ICU performance. Crit Care Clin 1996;12:503-14.
4. 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;9:591-7.
5. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: A severity of disease classification system. Crit Care Med 1985;13:818-29.
6. 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;34:1297-310.
7. Gupta R, Arora VK. Performance evaluation of APACHE II score for an Indian patient with respiratory problems. Indian J Med Res 2004;119:273-82.
8. Faruq MO, Mahmud MR, Begum T, Ahsan TS, 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;1:27-32.
9. Ayazoglu TA. A comparison APACHE II and APACHE IV scoring systems in predicting outcome in patients admitted with stroke to an intensive care unit. Anaesth Pain Intensive Care 2011;15:7-12.
10. Yamin S, Vaswani AK, Afreedi M. Predictive efficacy of APACHE IV at ICUs' of civil hospital Karachi. Pak J Chest Med 2011;17:11-21.
11. 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;142:851-8.
12. Kramer A, Zimmerman J, Jing Y. Performance of MPMo-III and APACHE IV mortality predictions in a multi-institutional ICU database: Implications for National Benchmarking. Crit Care Med 2012;40:1-328.
13. Brinkman S, Bakhshi-Raiez F, Abu-Hanna A, de Jonge E, Bosman RJ, Peelen L, *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;26:105.e11-8.
14. Kamal M, Khan NA, Ali G. A comparison of APACHE II and APACHE IV scoring systems in predicting outcome in patients with acute lung injury (ALI) and the adult respiratory distress syndrome (ARDS) in intensive care unit (ICU). RMJ 2013;38:234-8.
15. 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;67:115-22.
16. Costa e Silva VT, de Castro I, Liaño F, Muriel A, Rodríguez-Palomares JR, Yu L. 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;26:3894-901.
17. Tunnell RD, Millar BW, Smith GB. The effect of lead time bias on

severity of illness scoring, mortality prediction and standardised mortality ratio in intensive care – a pilot study. Anaesthesia 1998;53:1045-53.
18. Dünser MW, Baelani I, Ganbold L. A review and analysis of intensive care medicine in the least developed countries. Crit Care Med 2006;34:1234-42.

How to cite this article: 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;19:87-91.

Source of Support: Nil, **Conflict of Interest:** None declared.