Print ISSN: 2321-6379
Online ISSN: 2321-595X

Original Article

Comparison of Acute Physiology and Chronic Health Evaluation II and Acute Physiology and Chronic Health Evaluation IV Scoring System in Predicting Outcome in Trauma Patients Admitted in M.Y. Hospital, Indore

Sumit Shukla[1], Ankur Maheshwari[2], Deepak Shukla[3], Tanuj Ahirwar[3], Manishankar Das[4], Lakhan Parmar[4]

[1]Professor, Department of General Surgery, MGM Medical College, Indore, Madhya Pradesh, India, [2]Associate Professor, Department of General Surgery, MGM Medical College, Indore, Madhya Pradesh, India, [3]Post-Graduate III Year Resident, MGM Medical College, Indore, Madhya Pradesh, India

Abstract

Background: Acute physiology and chronic health evaluation (APACHE) scoring systems provide an objective means of mortality prediction in intensive care unit (ICU). The aims of this study were to compare the performance of APACHE II and APACHE IV in predicting mortality in our ICU.

Materials and Methods: A prospective observational study was conducted in a 25-bedded ICU of a tertiary level teaching hospital. All polytrauma patients between 18 and 36 years of age group, managed in ICU for >24 h, were enrolled. 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 up, and outcome was recorded as survivors or non-survivors. Observed mortality rates were compared with predicted mortality rates for both the APACHE II and APACHE IV. Receiver operator characteristic curves (ROC) were used to compare accuracy of the two scores.

Results: APACHE II score of the patients ranged from 1 to 30 and APACHE IV score of the patients ranged from 1 to 130. There was good correlation between APACHE II and APACHE IV scores with predictive mortality using Apache II 4.07 ± 5.83 and predictive mortality using Apache IV 7.87 ± 14.34. Discrimination for APACHE II and APACHE IV models was good with area under ROC curve of 0.987 and 0.994, respectively. APACHE IV was more accurate than APACHE II in this regard. The cutoff points with best Youden index association criterion for APACHE II were more than 12 and for APACHE IV were more than 54, respectively, for predicting mortality.

Conclusion: Discrimination was better for APACHE IV than APACHE II model; however, calibration was better for APACHE II than APACHE IV model in the present study. There was good correlation between the two models observed in the present study.

Key words: Acute physiology and chronic health evaluation II, Acute physiology and chronic health evaluation IV, Intensive care unit, Mortality prediction

INTRODUCTION

Intensive care unit (ICU) scoring systems were introduced almost 30 years ago with the goal of using physiologic data available at ICU admission to predict individual patient outcomes. The most commonly used scoring system in ICU-Acute Physiology and Chronic Health Evaluation II (APACHE II) was developed three decades back in 1985.[1]

APACHE scoring systems provide an objective means of mortality prediction in ICU.

The basic premise of these scores is that worst physiological derangement noted during the first 24 h after admission in

Access this article online

www.ijss-sn.com

Month of Submission : 04-2021
Month of Peer Review : 04-2021
Month of Acceptance : 05-2021
Month of Publishing : 06-2021

Corresponding Author: Deepak Shukla, M.G.M Boys Hostel Block-3, Indore, Madhya Pradesh, India.

an ICU determines the chance of hospital survival as these define organ insufficiency.

APACHE-I was proposed in the year 1981.[2] APACHE-I considered 34 routinely collected physiological measurements in the hospitals. Each of these measurements was assigned a weight according to the severity of derangement; higher the score more is the chance of death. However, this was found too complex as it included the large number of variables.[3]

APACHE II scoring system developed in 1985 as a modification of the original APACHE score, consists of reduced acute physiology score variables from 34 to 12 with age points and chronic health points.[1] The total physiological derangement score is the sum of the individual scores (0–4) for each variable, except the Glasgow Coma Scale (GCS) where the score is 15 minus the GCS (0–12).[4] The most deranged value in the first 24 h of ICU admission is used as the scoring for each variable. The total physiological derangement score is added to a score of age (0–6) and a chronic health score for patients with severe organ insufficiency (2–5 depending on admission status). A total score of 71 points is assigned which is used to relate to mortality in the ICU setting.

APACHE-III scoring system developed between 1991 and 1998 in several different versions. The APACHE III scores vary between 0 and 299 points, including up to 252 points for the 18 physiological variables, up to 24 points for age, and up to 23 points for the chronic health status.[5]

APACHE IV scoring system was introduced in 2006 as an improved and updated model for predicting hospital mortality among critically ill patients and is the most recent version of the APACHE scoring system. This model included the new predictor variables such as mechanical ventilation, thrombolysis, Pao2/fio2 ratio, impact of sedation on GCS, pre-ICU hospital length of stay, location before ICU admission, and 116 disease-specific subgroups in addition to the modifications introduced in the APACHE III.[6]

There are very few studies comparing the APACHE II and APACHE IV scoring systems in the ICU in an India setting.[7]

Thus, we designed a study to observe the performance of APACHE II and APACHE IV scoring system in our ICU. The newest scoring system from the APACHE foundation the APACHE IV attempted to improvise on the accuracy of outcome prediction.

MATERIALS AND METHODS

This was a prospective epidemiological study of polytrauma patients admitted at Trauma Centre, M.G.M. Medical College and M.Y. Hospital, Indore, Madhya Pradesh (M.P.), Indore, is a historic and main city in the state of M.P, India.

Study Period

This study period was from November 2018 to July 2020.

Sample Size

One hundred patents of polytrauma admitted at the Trauma Centre, M.G.M. Medical College and M.Y. Hospital, Indore.

Inclusion Criteria

The following criteria were included in the study:

1. All polytrauma patients between 18 and 36 years of age group.
2. Patients who give written informed consent.

Exclusion Criteria

The following criteria were excluded from the study:

1. Patient not willing to give written consent
2. Patients under the age of 18 years.
3. Patients above the age of 36 years.

Written consent was obtained from the patient or attendant. A comprehensive history was taken from the patient or the attendant and questionnaire filled for every patient. A set protocol was made for the management of polytrauma patient and patient was evaluated according to it.

Simple, rapidly performed maneuver's such as the administration of intravenous fluids, endotracheal intubation, and compressive dressings on sites of active emergency on arrival of patients to trauma center done.

Resuscitation primary priorities –

- Airway
- Breathing
- Circulation – Hemorrhage control
- Neurological examination
- Exposure of the patient

After resuscitation required investigations performed:

CBC count, coagulation studies, blood type, and blood cross-match (if indicated). Urinalysis, urine toxicological screen, and serum electrolyte values, creatinine level, and glucose values are often obtained for reference. Lipase or amylase level, Imaging studies (USG, X-RAY, CT Scan, MRI etc.).

On the day after ICU admission, the worst values on APACHE IV and APACHE II variables (worst

measurements observed during 24 h following ICU admission) were abstracted from clinical and laboratory records and APACHE scores were calculated using an online APACHE score calculator. Observed mortality rates were compared with predicted mortality rates for both the scoring systems and standardized mortality ratios (SMR) and sensitivity and specificity were determined. APACHE-IV predicted ICU patients were compared with observed ICU LOS and days on mechanical ventilation. Statistical analysis was carried out using a software package (SPSS for Windows; version 20.0) and $P < 0.05$ was considered statistically significant.

RESULTS

There were 15.0% of patients in the age group of 18–20 years, 40.0% of patients were in the age group of 21–30 years, and 45.0% of patients were in the age group of 31–36 years. Male patients were relatively younger with a mean age of 28.835 ± 5.829 years compared to 31.6 ± 5.527 years for female. In 24.0% of patients, the diagnosis was blunt trauma abdomen, in 32.0% of patients, it was penetrating injury, and in 44.0% of patients, the diagnosis was polytrauma. About 50% of patients were managed conservatively. In 50% of patient's, surgical management was done, in 44.0% of patients, exploratory laparotomy was done, in 3.0% of patients, craniectomy was done, in 3.0% of patients, craniotomy was done. In 19.0% of patients, there was requirement of mechanical ventilation and in 81.0% patients there was no requirement of mechanical ventilation [Table 1]. Total 100 polytrauma patients were included in my study in which 16 (16%) died and 84 (84%) were discharged [Table 2]. The poor calibration of APACHE II compared to APACHE IV in our study is also reflected by the SMR calculated using these scores. The SMR as per APACHE II was 0.84 and as per APACHE IV was 0.94. The case distribution of APACHE II and APACHE IV scoring system is discussed in Tables 3 and 4 respectively.

Table 1: Outcome data

	Patients no.	Mean days	Survivor	Non- survivor
ICU LOS	100	8.17±4.441	84	16
Ventilator days	19 out of 100	3.211±1.96		
Ventilator-free days of survivor	5 out of 19	9.4± 2.074	5	14

Table 2: Final outcome

Final outcome	Number	Percentage
Death	16	16.0
Discharge	84	84.0
Total	100	100.0

RECEIVER OPERATOR CHARACTERISTIC (ROC) CURVE ANALYSIS

Association between APACHE II and APACHE IV Using the Cutoff Generated by ROC Curves

ICU scoring systems were introduced many years ago with the goal of using physiologic data available at ICU admission to predict individual patient outcomes. Although these predictions have little utility for managing individual patients, they provide a mechanism to assess ICU performance by comparing actual outcomes in a given population to the outcomes observed in the reference population used to develop the prediction algorithms.[8]

Male patients were relatively younger with a mean age of 28.835 ± 5.829 years compared to 31.6 ± 5.527 years for female.

The mean APACHE II scores in our patients were 6.16 ± 7.07 (mean) [Table 5]. The non-survivors had a higher APACHE II score compared to survivors (3.583 vs. 19.688). APACHE II scores observed in our study were comparable to the values obtained from other studies by Parajuli *et al.*[9] (mean ± SD) score of 18.26 ± 7.4, 16.39 ± 6.82, and 22.08 ± 7.18 (survivors vs. non-survivors) and by Lee *et al.*[10] (mean ± SD) score of 16.9 ± 6.8 with 16.6 ± 6.6 for survivors and 26.1 ± 6.9 for non-survivors. The higher APACHE II scores indicate the increasing acuity of illness in the unit over the past decade.[11]

The mean APACHE IV score in our patients was 42.79 ± 24.77 (mean ± SD) [Table 5]. Our study results of

Table 3: Case distribution as per APACHE II scores

APACHE II score	No. of patients
0–5	66
6–10	13
11–15	8
16–20	4
21–25	7
26–30	2

Table 4: Case distribution as per APACHE IV scores

APACHE IV score	No. of patients
0–20	3
21–30	36
31–40	29
41–50	11
51–60	5
61–70	4
71–80	2
81–90	2
91–100	2
101–110	3
111–120	2
121–130	1

APACHE IV scores among survivors and non-survivors (mean) [Table 6] (33.643 vs. 90.812) were similar to APACHE IV scores reported by Lee *et al.*[10] (mean ± SD) (49 ± 22.2 vs. 77.1 ± 22.2) (survivors vs. non-survivors). However, Ayazoglu[12] reported a higher APACHE IV scores among survivors and non-survivors (mean ± SD) (79.9 ± 11.6 vs. 105.4 ± 14.9). Our ICU population is of polytrauma patients (surgical) and those of Lee *et al.*[10] were also exclusively surgical while of Ayazoglu were patients with stroke (medical). This difference in case mix explains the difference in scores observed in these units.

The discrimination of APACHE II as determined by AUC in our study was 0.987 (0.941–0.999: 95% confidence interval [CI]) [Table 7], which was almost similar to the AUC observed by Parajuli *et al.*[9] (AUC – 0.73). Discrimination of scoring systems changes when it is applied on different patient populations. This explains the varying AUC observed in the previous studies.[10,12,13] Our study population contained polytrauma surgical patients. Ayazoglu[12] had AUC – 0.98 on stroke patients while Lee *et al.*[10] and Brinkman *et al.*[13] obtained and AUC of 0.85 and 0.84 on post-operative patients and surgical patients, respectively. Discrimination of a scoring system also changes overtime.

The AUC for APACHE IV in our study was found to be 0.994 (0.953–1.000: 95% CI) [Table 8] indicating a better discriminating ability compared to APACHE II. Similar results were observed by Parajuli *et al.*[9] and Brinkman *et al.*[13] However, studies by Lee *et al.*[10] and Ayazoglu[12] found that APACHE II had a better discriminative power compared to APACHE IV [Table 9] in the original internal validation study for APACHE IV done by Zimmerman *et al.*[2] in the USA population, the discrimination was found to be very good with AUC of 0.880.

Table 5: Mean and standard deviations of various parameters

Parameter	No.	Mean±SD	Range
APACHE II score	100	6.16±7.07	0–26
Predictive mortality using APACHE II	100	4.07±5.83	0.5–27.9
apache iv score	100	42.79±24.77	18–125
Predictive mortality using APACHE IV	100	7.87±14.34	0.3–72.2

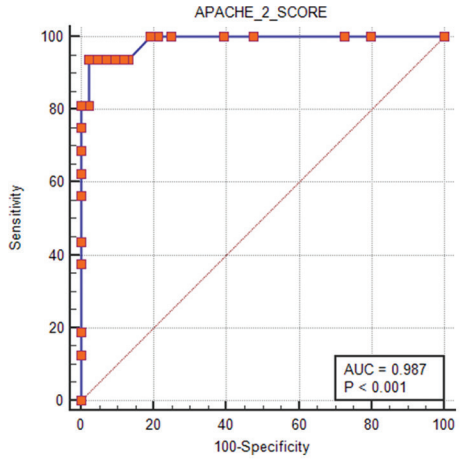
Table 6: Difference in APACHE II and IV scores between survivors and non-survivors

Survivor/Non-survivor	Mean values
	Mean values of APACHE II score
Survivor	3.583±3.581
Non-survivor	19.688±5.16
	Mean values of APACHE IV score
Survivor	33.643±10.96
Non-survivor	90.812±21.361

The AUC is 0.987 which shows that APACHE II is having outstanding discrimination for diagnosis of dead/discharge in patients with trauma. The cutoff value obtained is >12. If the APACHE II score is more than 12, then the likelihood of the patient dying is very high and if the score is ≤12, the likelihood of being discharged is very high. At this cutoff value, the sensitivity of APACHE II is 93.75% and specificity is 97.62%. The study by Ayazoglu *et al.* had AUC of 0.98 which was higher than all other reported literature.[12] This was likely due to the homogenous population of only stroke patients included in their study.

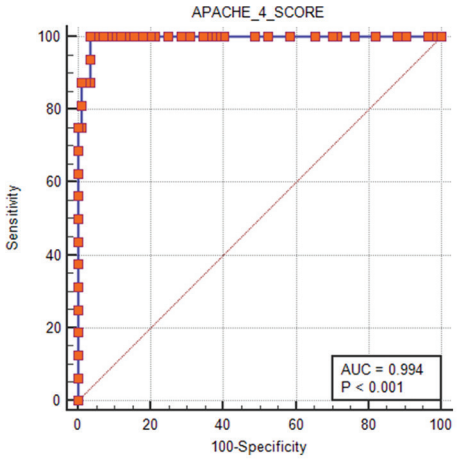
The AUC is 0.994 which shows that APACHE IV is having outstanding discrimination for diagnosis of dead/discharge in patients with trauma. The cutoff value obtained is >54. If the APACHE IV score is more than 54, then the likelihood of the patient dying is very high and if the score is ≤54, the likelihood of being discharged is very high. At

Table 7: ROC between APACHE II score and actual outcome (dead/discharge)



Variable	APACHE_2_SCORE
Classification variable	OUTCOME_NO
Sample size	100
Positive group[a]	16 (16.00%)
Negative group[b]	84 (84.00%)
[a]OUTCOME_NO = 1	
[b]OUTCOME_NO = 0	
Disease prevalence (%)	Unknown
Area under the ROC curve (AUC)	
Area under the ROC curve (AUC)	0.987
Standard error[a]	0.0107
95% confidence interval[b]	0.941–0.999
Z statistic	45.445
Significance level *P* (area=0.5)	⟨<0.0001⟩
[a]DeLong *et al.*, 1988	
[b]Binomial exact	
Youden index	
Youden index J	0.9137
Associated criterion	>12
Sensitivity	93.75
Specificity	97.62

Table 8: ROC between APACHE IV score and actual outcome (dead/discharge)



Variable	APACHE_4_SCORE
Classification variable	OUTCOME_NO
Sample size	100
Positive group[a]	16 (16.00%)
Negative group[b]	84 (84.00%)
[a]OUTCOME_NO=1	
[b]OUTCOME_NO=0	
Disease prevalence (%)	Unknown
Area under the ROC curve (AUC)	
Area under the ROC curve (AUC)	0.994
Standard Error[a]	0.00474
95% Confidence interval[b]	0.953–1.000
Z statistic	104.229
Significance level *P* (area=0.5)	⁅<0.0001⁆
[a]DeLong *et al*., 1988	
[b]Binomial exact	
Youden index	
Youden index J	0.9643
Associated criterion	>54
Sensitivity	100.00
Specificity	96.43

this cutoff value, the sensitivity of APACHE IV is 100.00% and specificity is 96.43%. Result of study for APACHE IV done by Zimmerman *et al.* with AUC of 0.880.[2]

We observed that the APACHE IV model better predicted mortality than APACHE II scoring system in our ICU. The reason could probably be the consideration of mechanical ventilation support, patient source before ICU admission, and specific reason for ICU admission included in the APACHE IV model.

In our study, the cutoff point with best Youden index for APACHE II was 12 and APACHE IV was 54. Hence, mortality of patients was significantly high when APACHE II score was ≥12 and APACHE IV score ≥54 [Tables 10 and 11].

The poor calibration of APACHE II compared to APACHE IV in our study is also reflected by the SMR calculated using these scores. The SMR as per APACHE II was 0.84 and as per APACHE IV was 0.94. Both the

Table 9: Comparison of discrimination between APACHE II and APACHE IV in various studies

	APACHE II	APACHE IV
Our study	0.987	0.994
Parajuli *et al.*	0.73	0.79
Brinkman *et al.*	0.84	0.87
Lee *et al.*	0.85	0.8
Ayazoglu *et al.*	0.98	0.93

Table 10: Sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy of APACHE II against APACHE IV score in the diagnosis of dead/discharge (considering here APACHE IV as the gold standard)

APACHE-II	APACHE-IV		
	Dead (>54)	Discharge (≤54)	Total
Dead (>12)	17	0	17
Discharge (≤12)	2	81	83
Total	19	81	100

The above table shows the sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy of APACHE II against APACHE IV.
For this analysis, we have considered APACHE IV as the gold standard.
Sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy of APACHE II is:

Sensitivity	89.47%
Specificity	100.00%
Positive predictive value	100.00%
Negative predictive value	97.59%
Diagnostic accuracy	98.00%

Table 11: Sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy of APACHE IV against APACHE II score in the diagnosis of dead/discharge (considering here APACHE II as the gold standard)

APACHE-IV	APACHE-II		
	Dead (>12)	Discharge (≤12)	Total
Dead (>54)	17	2	19
Discharge (≤54)	0	81	81
Total	17	83	100

The above table shows the sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy of APACHE IV against APACHE II.
For this analysis, we have considered APACHE II as the gold standard.
Sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy of APACHE IV is:

Sensitivity	100.00%
Specificity	97.59%
Positive predictive value	89.47%
Negative predictive value	100.00%
Diagnostic accuracy	98.00%

predictive scoring systems overestimated the overall mortality rate. The SMR as per APACHE II was 0.72 in the earlier study by Arunkumar *et al*.[11] Similarly, poor calibration observed by Lee *et al*.[10] was also reflected in the very low SMR of 0.11 with APACHE II and 0.21 with APACHE IV in their study.

The limitations for these scoring systems would be failure to compensate for lead time bias, requirement for a diagnosis immediately on admission, pre-ICU conditions and management, and poor interobserver reliability. Besides this, overall quality of ICU is affected by the bed occupancy ratio, laboratory facility, availability of trained workforce, nurse to patient ratio, and financial status of the patient's caregiver. Both APACHE models have shown good performance in the developed countries and other parts of the world.

Temporal bias is another factor responsible for the poor calibration in external validation studies of scoring systems. The APACHE II scoring system was developed three decades back and the APACHE IV almost a decade back. Medical science and quality of ICU care has improved exponentially in the meantime causing the predictive scoring systems to become less accurate. The differences in sample size between the study population and the original cohort used in the development of the scoring systems can also lead to a difference in the predictive accuracy.

Hence, predictive scoring systems developed in a western population needs customization before they can be applied to Indian population.

SUMMARY AND CONCLUSION

Our study is not short of limitations. This was a single-center study over as short span of time with relatively fewer numbers of patients compared to the original internal validation study. We did not look into the disease-specific performance of APACHE IV. We have also not analyzed the lead time of our patients to ICU admission, which is an important factor that can impact the performance of scoring systems.

In summary, the APACHE II and APACHE IV scoring systems showed reasonably good discriminating ability in our ICU though both calibrated poorly. APACHE II showed better calibration than APACHE IV. Larger multicenter validation studies with customization for the Indian ICU population are needed before they can be applied in our setting.

In our ICU, APACHE IV model showed better predicted mortality than APACHE II scoring system. Moreover, calibration of both the scoring systems was less than adequate. However, APACHE II showed better calibration than APACHE IV. Larger multicentered validation studies with customization for the Indian ICU population are needed before they can be applied in our setting.

ACKNOWLEDGMENT

Authors acknowledge the enormous help received from the authors whose articles are cited and included in references to this manuscript. The authors are also grateful to authors/editors/publishers of all those articles, journals, and books from where the literature for this article has been reviewed and discussed.

REFERENCES

1. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: A severity of disease classification system. Crit Care Med 1985;13:818-29.
2. 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.
3. Wong DT, Knaus WA. Predicting outcome in critical care: The current status of the APACHE prognostic scoring system. Can J Anaesth 1991;38:374-83.
4. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: A practical scale. Lancet 1974;304:81-4.
5. Zimmerman JE, Wagner DP, Draper EA, Wright L, Alzola C, Knaus WA. Evaluation of acute physiology and chronic health evaluation III predictions of hospital mortality in an independent database. Crit Care Med 1998;26:1317-26.
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. Breslow MJ, Badawi O. Severity scoring in the critically Ill. Chest 2012;141:245-52.
9. 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.
10. 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.
11. Arunkumar AS, Rajaram MU, Kamat V. Validation of APACHE II and SAPS II in a multi-disciplinary ICU. Indian J Crit Care Med 2002;6:12-8.
12. Ayazoglu TA. Validation of the APACHE IV scoring system in patients with stroke. A comparison with the APACHE II system. Anaesth Pain Intensive Care 2011;15:7-12.
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.

How to cite this article: Shukla S, Maheshwari A, Shukla D, Ahirwar T, SDas M, Parmar L. Comparison of Acute Physiology and Chronic Health Evaluation II and Acute Physiology and Chronic Health Evaluation IV Scoring System in Predicting Outcome in Trauma Patients Admitted in M.Y. Hospital, Indore. Int J Sci Stud 2021;9(3):65-70.

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