East African Medical Journal Vol. 102. 7 July 2025
EVALUATING THE PERFORMANCE OF APACHE II AMONG CRITICALLY ILL PATIENTS ADMITTED AT MOI TEACHING AND REFERRAL HOSPITAL, ELDORET – KENYA
Edwin Gudu, Moi Teaching and Referral Hospital (MTRH), P.O. Box 3, Eldoret, Kenya, email: edwinosenogudu@gmail.com, Registrar, Department of Surgery and Anaesthesiology, Kituyi Werunga, Consultant Anaesthesiologist, Department of Surgery and Anaesthesiology, Moi University – School of Medicine, P.O. Box 4606, Eldoret, Kenya, email: kituyipw@gmail.com, Elisha Kirwa, Moi Teaching and Referral Hospital (MTRH), P.O. Box 3, Eldoret, Kenya, email:kirwaz@gmail.com, Consultant Anaesthesiologist, Department of Surgery and Anaesthesiology, Cecilia Kiilu, Consultant Paediatrician, Department of Child Health and Paediatrics, Moi University – School of Medicine, P.O. Box 4606, Eldoret, Kenya, email: cckiilu@hotmail.com

Corresponding author: Edwin Gudu, Department of Surgery and Anaesthesiology, Moi Teaching and Referral Hospital (MTRH), P.O. Box 3, Eldoret, Kenya, email, edwinosenogudu@gmail.com

EVALUATING THE PERFORMANCE OF APACHE II AMONG CRITICALLY ILL PATIENTS ADMITTED AT MOI TEACHING AND REFERRAL HOSPITAL, ELDORET - KENYA

E. Gudu, K. Werunga, E. Kirwa, C. Kiilu

ABSTRACT

Objective: **This study evaluated the performance of the acute physiology and chronic health evaluation (APACHE) II model among critically ill patients admitted at Moi teaching and referral hospital (MTRH) intensive care unit (ICU) by assessing its calibration and discriminative ability.**
Design: The study was a prospective analytical cross-sectional study
Setting: The study was carried out at MTRH, a public tertiary teaching hospital in Eldoret, Kenya. It has a 20-bed mixed use ICU admitting both surgical and medical patients.
Participants: **A total of 416 participants aged over 18 years were consecutively recruited.**
Main outcome measures: **For the categorical variables, odds ratios were calculated at 95% confidence interval and p-values <0.05 considered statistically significant. Area under receiver operating characteristics (AUROC) curve was used to assess the discriminative ability of the model to distinguish between survivors and non-survivors. Hosmer-Lemeshow chi square statistic was used to assess the goodness of fit of the model.**
Results: **Of the 416 participants, there was a 57% (n=237) male majority. The ICU mortality rate was at 29% (n=120). Fever (AOR 1.2, 95% CI 0.1 – 13.7), low mean arterial pressure (AOR 1.8, 95% CI 0.5 – 6.1), tachycardia (AOR 1.9, 95% CI 0.9 – 4.1), hypoxaemia (AOR 1.1, 95% CI 0.5 – 2.5), acidosis (AOR 3.2, 95% CI 1.5 – 6.8) and hyponatremia (AOR 1.2, 95% CI 0.4 – 3.2) were associated with higher odds mortality. The AUROC curve yield for the model was 0.77 whilst the p-value for the Hosmer-Lemeshow statistic was 0.03, which showed a poor calibration.**
Conclusion: **The APACHE II has acceptable accuracy in discrimination of survivors and non-survivors and can be used as a clinical adjunct in decision-**

making in critically ill patients. Additional studies in resource limited settings required in order to improve its calibration.

INTRODUCTION:

Intensive care units are vital in the management of critically ill patients. The demand for intensive care unit (ICU) services will always far outstrip the available intensive care unit capacity especially in resource constrained settings. Kenya, just like other developing countries, has a shortage of available ICU beds accounting for only 0.9% of the total number of hospital beds (1). This is against a Society of Critical Care Medicine recommendation of 20 - 40 % ICU beds. As such, proper resource allocation and utilization, is vital to ensure that the available intensive care capacity achieves the maximum clinical impact.

Several outcome prediction models have been developed to score disease severity, prognosticate and predict outcomes among populations of critically ill patients. The models use various predictor variables such as: acute physiological state, clinical information and laboratory findings to determine the likelihood of a clinical outcome. These models can be used as clinical adjuncts in determining the persons most likely to benefit from the scarce intensive care unit services. Despite the validation of outcome predictive models for use in resource limited settings, their uptake and utilization is still low.

One such model is the Acute physiology and chronic health evaluation (APACHE II). It uses clinical and laboratory variables that are usually available at ICU admission. It has been validated for use in both high and low resource settings. This study aimed to evaluate its utility as clinical adjunct in determination of persons most likely to benefit from ICU admission. This was achieved by evaluating the model's variables as well as the calibration and discriminative ability.

METHODS

The study was an analytical cross-sectional study carried out at Moi Teaching and Referral Hospital (MTRH) intensive care unit, a tertiary teaching hospital, located in Eldoret, Kenya. The study was carried out between October 2022 and July 2023. MTRH intensive care unit is a 20-bed mixed-use medical and surgical unit attending to adult and paediatric patients.

The minimum sample size was calculated using the Cochran formula and sensitivity of the APACHE II model in low resource settings at 88% (2). This yielded 163 persons. However, given that this was a sub-study under a main study that was comparing the APACHE II model to two other outcome predictive models, the overall sample size used was 416 persons same as that in the main study. The participants were recruited consecutively. Persons aged less than 16 years were excluded since this model is not recommended for use in this age group. This study, being prospective in nature, also excluded those persons who had missing data on the variables of interest that were required in the calculation of the mortality probability percentage.

Clinical outcomes (survivors versus non-survivors) and length of intensive care unit admission were the dependent variables of interest. The independent variables used during this study included: demographic data such as age and gender; clinical data such as the date of admission, the ward of origin prior to ICU admission, co-morbidities present, working diagnosis, date of discharge and clinical outcome; laboratory findings for tests such as full haemogram, renal function tests and blood gas analysis; vital signs and physiological state of the participants at ICU admission; the need for mechanical ventilation and finally the need for inotropic or vasopressor support.

The data analysis was carried out using R® statistical software version 2023.09.0 + 463. This study used area under receiver operating characteristics curve (AUROC) to assess the discriminatory ability of the model to differentiate between the survivors and non-survivors. In addition, the AUROC curve was also used to calculate the sensitivity and specificity of the model. Hosmer-Lemeshow chi square statistic was used to assess the calibration of the model by comparing the estimated non-survivors with the actual numbers of the non-survivors. For this test, if the p-value for the chi square statistic was > 0.05, the model was considered to be having a good calibration. This shows that there is no statistically significant difference between the predicted outcomes and the actual observed outcomes. For continuous variables, frequencies, proportions, measures of central tendency and dispersion were calculated. For categorical variables, odds ratios were calculated at 95% confidence level and p-values <0.05 considered to be statistically significant.

All ethical considerations were taken into account prior to the onset of the study. Ethical approval was obtained from the Moi University / MTRH institutional research and ethics committee (formal approval number 0004097) and a research permit obtained from the National Commission for Science, Technology and Innovation (NACOSTI) - Research license number 608027 prior to the onset of the study. Permission was also obtained from the MTRH management prior to commencement of the data collection. Written informed consent was also obtained from the study participants prior to enrollment in the study. Confidentiality was maintained throughout the study and the data de-identified prior to analysis.

RESULTS

A total of 416 participants were recruited. There was a 57% (n=237) male majority. The average age of the study participants was 42 years (SD 20). During the study period, the ICU mortality rate was 29%. The average age of the non-survivors was 45 years (SD 21). The median length of ICU stay among the non-survivors was 3 days (IQR 1, 7 days).

The sensitivity of the model was 58% with a specificity of 84%. The area under the receiver operating characteristics curve (AUROC) yield was 0.77 as shown in figure 1 below:

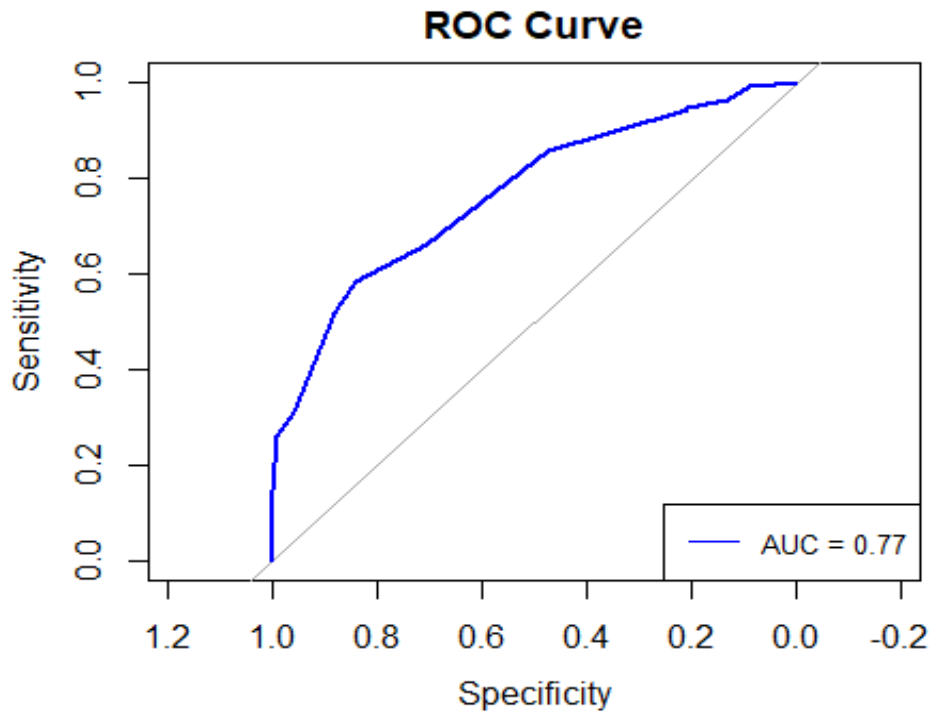

Figure 1:
Area under receiver operating characteristics curve for the APACHE II Model

The Hosmer-Lemeshow chi square statistic showing the expected (predicted) against the actual (observed) mortalities was 13.6 with a p value of 0.03 showing that the model had poor calibration. This is as shown in table 1 below:

Table 2:
Distribution of expected and observed outcomes against estimated mortality probabilities using APACHE II

		Non-survivor		Survivor	
Probability of mortality	**Total**	**Observed**	**Expected**	**Observed**	**Expected**
0 - 10	97	5	5.238	92	91.762
11 - 20	95	14	12.160	81	82.840
21 - 30	39	16	11.076	23	27.924
31 - 40	110	42	41.580	68	68.420
41 - 50	34	18	14.960	16	19.040
51 - 60	14	9	7.700	5	6.300
71 - 80	13	5	9.490	8	3.510
81 - 90	14	11	12.110	3	1.890
Total	**416**	**120**		**296**	

Several variables from the APACHE II model showed statistically significant associations with increased mortality. Of note, the acute physiological variables had high mortality predictive ability. This included vital signs such as body temperature, blood pressure and heart rate. Elevated body temperature greater than 38.4°C was associated with a fourfold increase in mortality risk (odds ratio [OR] 4.0; 95% confidence interval [CI], 1.1–44.5; p = 0.02). A mean arterial pressure below 70 mmHg also correlated with higher mortality (OR 2.0; 95% CI, 1.1–3.6; p = 0.01). Tachycardia, defined as a heart rate exceeding 109 beats per minute, was significantly associated with increased mortality risk (OR 2.8; 95% CI, 1.8–4.5; p < 0.001).

The APACHE II model also requires laboratory tests for its estimation of mortality probability. Full haemogram, blood gas analysis and renal function tests have to be done. On blood gas analysis, hypoxaemia, indicated by a partial arterial pressure of oxygen (PaO_2) less than 70 mmHg, showed an increased association with mortality (OR 1.9; 95% CI, 1.2–3.0; p = 0.003). Having acidosis, pH < 7.33, at the time of ICU admission had a similar association (OR 2.4; 95% CI, 1.5–3.8; p < 0.001).

Electrolyte disturbances, including hyperkalemia (>5.4 mmol/L) and hyponatremia (<130 mmol/L), were also significant predictors of mortality (OR 2.5; 95% CI, 1.2–5.0; p = 0.009 and OR 2.9; 95% CI, 1.7–4.9; p < 0.001, respectively). The other renal derangement evidenced by elevated serum creatinine (>124 μmol/L) at the time of ICU admission also showed an increased mortality risk (OR 2.5; 95% CI, 3.8–6.5; p < 0.001). On full haemogram, both leucopenia (white blood cell count <3.0 × 10^3/μL; χ^2 = 11.4; p < 0.001) and leucocytosis (>14.9 × 10^3/μL; OR 1.7; 95% CI, 1.1–2.7; p = 0.03) were positively associated with higher mortality risk.

This summary is as shown in table 2 below:

Table 2:

Table showing summary of the performance of the variables used in the APACHE model

Variable		Outcome		OR (95% CI)	P-value	AOR (95% CI)	P-value
	Ranges	Non - survivor	Survivor				
Temperature *(^0C)*	36.0 - 38.4 *(Referent group)*	32	245				
	Low (<36.0)	22	47	1.2 (0.7 – 2.2)	0.44		
	High (>38.4)	6	4	4.0 (1.1 – 44.5)	0.02	1.3 (0.1 – 13.9)	0.85
Blood pressure *(Mean arterial pressure in mmHg)*	70 – 109 *(Referent group)*	80	222				
	Low (<70)	25	34	2.0 (1.1 – 3.6)	0.01	1.8 (0.5 – 6.3)	0.33
	High (>109)	14	40	1.0 (0.5 – 2.0)	0.9		

Heart rate *(Beats per minute)*	70 - 109 *(Referent group)*	51	188				
	Low (<70)	10	31	1.2 (0.5 – 2.6)	0.66		
	High (>109)	59	77	2.8 (1.8 – 4.5)	<0.001	2.0 (1.0 – 4.3)	0.08
Respiratory rate *(Breathes per minute)*	12 - 24 *(Referent group)*	94	243				
	Low (<12)	7	16	1.1 (0.5 – 2.8)	0.79		
	High (>24)	19	37	1.3 (0.7 – 2.4)	0.36		
Oxygenation *($Pa0_2$ in mmHg)*	Normoxaemic (≥70)	72	220	1.9 (1.2 – 3.0)	0.003	1.2 (0.6 -2.8)	0.60
	Hypoxaemic (<70)	48	76				
Serum PH on blood gas analysis	7.33 - 7.49 *(Referent group)*	54	182				
	Low (<7.33)	57	79	2.4 (1.5 – 3.8)	<0.001	3.3 (1.5 -7.2)	0.003
	High (>7.49)	9	35	0.9 (0.4 – 1.9)	0.72		
Serum sodium levels *(mmol/L)*	130 - 149 *(Referent group)*	76	244				
	Low (<130)	36	40	2.9 (1.7 – 4.9)	<0.001	1.1 (0.4 – 3.0)	0.87
	High (>149)	8	12	2.1 (0.8 – 5.4)	0.1		
Serum potassium levels *(mmol/L)*	3.5 - 5.4 *(Referent group)*	84	232				
	Low (<3.5)	19	45	1.2 (0.6 – 2.1)	0.61		
	High (>5.4)	17	19	2.5 (1.2 – 5.0)	0.009	0.3 (0.1 – 1.5)	0.15
Serum creatinine levels *(micromol/L)*	53 - 124 *(Referrent group)*	60	202				
	Low (<53)	19	58	1.1 (0.6 – 2.0)	0.75		
	High (>124)	41	36	3.8 (2.3 – 6.5)	<0.001	1.1 (0.4 – 3.3)	0.80
White Blood Cell Count (counts per microlitre)	3,000 - 14,900 *(Referrent group)*	75	223				
	Low (<3,000)*	4	0				
	High (>14,900)	41	73	1.7 (1.1 – 2.7)	0.03	1.3 (0.6 – 3.0)	0.47

DISCUSSION

In this study, based on the AUROC curve, the APACHE II performance was "good". This was scored using the grading criteria chosen for this study (2). This shows adequate discriminative power of the tool. As such, it has reasonable predictive power to distinguish between the survivors and non-survivors among the patients being admitted into the intensive care unit. This good performance can be partly attributed to the performance of the individual variables used in this model. Of note is that the variables, especially the acute physiological and laboratory derangements, had significant mortality predictive ability. This is because each of the variables look at the various body organ systems in order to assess and give a disease severity score, and by extension, the mortality probability estimate (3,4).

The APACHE II model has been tested in both high and low resource countries and many of the studies have shown a similar trend in terms of the discriminative ability of the tool. The AUROC curves have shown yields of between 0.70 to 0.86 corresponding with results of "good" and "very good" on the scoring table. Most of the studies from countries other than the United States of America (USA), from where the initial APACHE II model was developed, scored "good" based on the AUROC curve values (3,5–9).

This shows the importance of customizing the APACHE II model for suitable use in the country in which it is intended to be deployed. This is done by customizing and generating beta coefficients for use in the APACHE II logistic regression models from local studies intending to use it (10). For example, additional customization has been done on the APACHE II model. In 2006, the APACHE IV model, the most current of the APACHE series of models, was validated for use in the United States of America (USA) after a large clinical trial restricted to the critical care population/patients in hospitals across the USA. Subsequent tests carried out showed better discriminative ability of APACHE IV as compared to the APACHE II (10–13).

Calibration of a mortality prediction model refers to its measurement / the agreement of expected and observed outcomes (survivors and non-survivors) along all the predicted range of probabilities. For example, there needs to be an agreement between the expected and observed values at predicted mortality of 10%, 60% and 90% and so forth (10). For this study, the calibration of the APACHE II was low as shown by the statistically significant p-value for the Hosmer-Lemeshow statistic (2,4). This can be interpreted as there being a statistically significant difference between the observed and the predicted outcomes. This can still be explained by the fact that calibration of mortality predictive models performs best in the population from which the beta coefficients for the logistic regression models was derived (10,14).

For this study, most of the variables used in APACHE II were found to be positively associated with increased mortality and were also mostly statistically significant on bivariate analysis. As such, the odds ratios derived from this study can be used to generate beta coefficients that can improve

both the discriminative ability as well as the calibration of the model.

LIMITATION

The main limitation to this study is that it was a single center study hence the case-mix of the study participants was limited to the patients admitted at the hospital.

CONCLUSION

Therefore, having taken into consideration the calibration and the discriminative ability of this scoring model, it can be deployed for use in our set up in situations where all the required variables are available.

RECOMMENDATION

Additional multicenter studies can be carried out in order to improve the calibration and discriminative ability of the model for deployment in our population and patient case mix.

CREDIT STATEMENT

Each author participated in the proposal development, research process, data analysis and manuscript writing

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare

REFERENCES

1. Barasa EW, Ouma PO, Okiro EA (2020) Assessing the hospital surge capacity of the Kenyan health system in the face of the COVID-19 pandemic. PLoS ONE 15(7): e0236308. https://doi.org/10.1371/journal.pone.0236308

2. Haniffa R, Isaam I, De Silva AP, Dondorp AM, De Keizer NF. Performance of critical care prognostic scoring systems in low and middle-income countries: A systematic review. Crit Care [Internet]. 2018 Jan 26 [cited 2021 Mar 27];22(1):18. Available from: https://ccforum.biomedcentral.com/articles/10.1186/s13054-017-1930-8

3. Mumtaz H, Ejaz MK, Tayyab M, Vohra LI, Sapkota S, Hasan M, Saqib M. APACHE scoring as an indicator of mortality rate in ICU patients: a cohort study. Ann Med Surg (Lond). 2023 Mar 24;85(3):416-421. doi: 10.1097/MS9.0000000000000264. PMID: 37008173; PMCID: PMC10060092.

4. Sungono V, Hariyanto H, Soesilo TEB, Adisasmita AC, Syarif S, Lukito AA, et al. Cohort study of the APACHE II score and mortality for different types of intensive care unit patients. Postgrad Med J [Internet]. 2022 Dec 1 [cited 2025 Jul 14];98(1166):914–8. Available from: https://dx.doi.org/10.1136/postgradmedj-2021-140376

5. Fernandes S, Sérvio R, Patrício P, Pereira C. Validation of the Acute Physiology and Chronic Health Evaluation (APACHE) II Score in COVID-19 Patients Admitted to the Intensive Care Unit in Times of Resource Scarcity. Cureus. 2023 Feb 7;15(2):e34721. doi: 10.7759/cureus.34721. PMID: 36909097; PMCID: PMC9998113.

6. Kostoglou A, Kotanidou A, Marvaki C, Orfanos S. Mortality rate of ICU patients according to APACHE II and SAPS II score. Health Res J. 2018 Dec 7;4(4):219-33. Available from: https://ejournals.epublishing.ekt.gr/index.ph

p/HealthResJ/article/view/19297 [cited 2025 Mar 2]

7. Ruiz de Gopegui Miguelena P, Martínez Lamazares MT, Claraco Vega LM, Gurpegui Puente M, González Almárcegui I, Gutiérrez Ibañes P, Carrillo López A, Castiella García CM, Miguelena Hycka J. Evaluating frailty may complement APACHE II in estimating mortality in elderly patients admitted to the ICU after digestive surgery. Med Intensiva (Engl Ed). 2022 May;46(5):239-47. doi: 10.1016/j.medine.2022.02.019. Epub 2022 Mar 2. PMID: 35248506.

8. Cao L, Cherukuri K, Vu T, Chinn J, Ahmadi T, Ferrante MDE, Chandak T. APACHE II score and mortality risk prediction in a US (Californian) cohort of critically ill patients with COVID-19. Chest. 2021 Oct;160(4):A1032. doi: 10.1016/j.chest.2021.07.957. Epub 2021 Oct 11. PMCID: PMC8503156.

9. Parker RK, Mwachiro EB, Mwachiro MM, Pletcher J, Parker AS, Many HR. Mortality Prediction in Rural Kenya: A Cohort Study of Mechanical Ventilation in Critically Ill Patients. Crit Care Explor [Internet]. 2019 Dec [cited 2021 Mar 26];1(12):e0067. Available from: https://journals.lww.com/10.1097/CCE.0000000000000067

10. Kelley MA. Predictive scoring systems in the intensive care unit - UpToDate [Internet]. Uptodate. 2020 [cited 2021 Mar 26]. Available from: https://www.uptodate.com/contents/predictive-scoring-systems-in-the-intensive-care-unit?search=mortality prediction models icu&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1#H941655083

11. Jeong S. Scoring systems for the patients of intensive care unit. Acute Crit Care. 2018 May;33(2):102-4. doi: 10.4266/acc.2018.00185. Epub 2018 May 31. PMID: 31723870; PMCID: PMC6849060.

12. Varghese YE, Kalaiselvan MS, Renuka MK, Arunkumar AS. Comparison of acute physiology and chronic health evaluation II (APACHE II) and acute physiology and chronic health evaluation IV (APACHE IV) severity of illness scoring systems, in a multidisciplinary ICU. J Anaesthesiol Clin Pharmacol. 2017 Apr-Jun;33(2):248-53. doi: 10.4103/0970-9185.209741. PMID: 28781454; PMCID: PMC5520601.

13. 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 May;7(5):1598-603. doi: 10.18203/2320-6012.ijrms20191643.

14. Hersh AM, Beesley SJ. The difficulty of predicting intensive care unit mortality in resource-limited settings. Ann Am Thorac Soc. 2018 Nov;15(11):1282-4. doi: 10.1513/AnnalsATS.201808-580ED. PMID: 30382783.