

Evaluation of APACHE II, RAPS, and REMS models at the emergency department to predict hospital death in triage levels 1 and 2

Hamed Shokoohsaremi*, Mohsen Ebrahimi, Azade Shirinzadeh Feizabadi

Department of Emergency Medicine, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad

RESEARCH

Please cite this paper as: Shokoohsaremi H, Ebrahimi M, Shirinzadeh Feizabadi A. Evaluation of APACHE II, RAPS, and REMS models at the emergency department to predict hospital death in triage levels 1 and 2. AMJ 2023;16(6): 647-651.

<https://doi.org/10.21767/AMJ.2023.3705>

Corresponding Author:

Hamed Shokoohsaremi

Student Research Committee, Department of Emergency Medicine, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

Email: h.saremi@aftermail.ir

ABSTRACT

Background

One of the essential components for many therapeutic decisions is the initial evaluation of the patients.

Aims

The initial evaluation is usually done by the physician, which is mainly performed subjectively and is estimated more accurately if the physician is experienced. Many systems have been developed to evaluate and make these evaluations objective. The present study has used three different scoring systems.

Methods

The present study was prospective cohort on the patients admitted with triage 1 and Two levels. The study was conducted over one year. APACHE II, RAPS, REMS models were calculated for each patient. Then these models were evaluated from three main aspects of overall performance, resolution and calibration.

Results

The present study evaluated 1029 patients admitted to Edalati Emergency Unit in Mashhad. From among them, 198 patients were at triage level 1 consistent with ESI algorithm and the rest were assigned to level two. Overall, about 29 percent of these patients died and 753 survived. The largest area below the curve was dedicated to APACHE II model, which shows the high discrimination of the model. AUC=0.76 (CI:0.72, 0.78). in the next ranks were REMS

model with the area of 0.67 and the RAPS model with the area of 0.63. Hosmer-Lemeshow test for all models was <0.05. In general, the performance of APACHE II models was more favourable.

Conclusion

The results showed that APACHE II model had a higher degree of differentiation compared to RAPS and REMS models. The only overall performance of APACHE II model was in a relatively reasonable range. However, overall the compliance of the predicted and actual mortality rates was low, and the need to recalibrate the models seems necessary.

Key Words

APACHE, RAPS, REMS, hospital

What this study adds:

1. What is known about this subject?

The role of predicting scoring systems in evaluating death in high risk patients enter the emergency department.

2. What new information is offered in this study?

APACHE2 is the most useable predicting system in emergency death performantly but its calibration rate is not acceptable.

3. What are the implications for research, policy, or practice?

Due to low range of calibrations, more studies are needed to achieve the best calibrating rates and make the systems accurated.

Background

Mortality prediction models are considered a familiar and long-term research component. These models provide a metric and standard measure for measuring the severity and intensity of the disease and evaluate death as a consequence. In addition, a variety of physiological scoring systems are designed for use in emergency designed to follow a common goal, which is to measure the extent of physiological deviations and subsequently determine the severity of the disease^{1,2}. These systems should have unique features, such as ease of use, clinical availability, accurate

prediction, short data collection time, and the ability to compare with ICU scoring systems.

Regarding this issue, the explanation of one or more emergency predictor systems that can measure the severity and intensity of a patient's condition and, consequently, the classification of services, will be reduced to a very efficient level from the rate of hospital mortality. The rapid diagnosis of ill patients and the implementation of vital measures in the shortest time possible are possible only with the prioritization of patients. Nowadays, due to the type of life and the increase in accidents, the importance of emergency units and their quality has been considered more than ever before. Among the changes with drastic effects on the performance of management systems in recent years is the existence of a system of evaluation and performance monitoring^{3,4}. The present study examined APACHE II, REMS, and RAPS models.

Methods

This study was conducted on 1029 patients as prospective. The study was conducted in one of the largest emergency centers in eastern Iran during 2016-2017. Patients over 18 enrolled in the study, of whom the ones who had admission history of less than four hours or died upon admission, trauma patients and poisoned patients were excluded from the study. Patients whose physiological points were not measurable because of lack of enough information and the ones with unknown identity were excluded from the study.

Discussion

Using scoring systems can diagnose the patients at risk, which may be ignored by the medical staff and improve their outcomes. In this study, APACHE II, RAPS, and REMS models were calculated for each patient.

The most sub-curve area is shifted to APACHE II model, which shows the high discrimination of the model. AUC = 0.76 (0.728-0.792) was the first and REMS model with 0.669 area was the next and RAPS model was in the next stage with 0.629. Moreover, as shown in Table 2, APACHE II model had the most desirable BS value. RAPS model underestimates all predicted death points for patients. In REMS model, which is a modified RAPS model, we witnessed relatively better predictor accuracy, although overall, this model underestimates death prediction. The calibration chart shows that the studied models did not have enough compliance with the patient's final outcome. Weakness in calibration is an important issue that should be considered before using practical models. Different causes can cause a difference in actual and predicted mortality. Even small differences in terms of composition are shown that affect score calibration⁵.

Additionally, patients who were re-enrolled, pregnant women and surgical cases were not included in the study. The comparison of the models was investigated considering three main dimensions: main function, power differentiation and calibration. The differentiation power of the model refers to diagnose the survival and death of patients using Roc curve graph. A binary comparison of ROC models was performed using the Delong's test technique. The Brier Score (BS) factor, calculated as a benchmark for prediction accuracy, was calculated according to the following formula for all models.

For model calibration, goodness of fit of Hosmer Lemeshow and the calibration curve were used. Moreover, R studio software was used for statistical analysis. In R program for the above analyses, the pROC statistical package and the Predictable statistical package were used.

Results

In this study, 1029 patients admitted to the emergency units were evaluated. From among them, 198 patients were at triage level 1 consistent with ESI algorithm and the rest were assigned to level two. Overall, about 29 percent of these patients died and 753 survived (413 men and 340 women) and 276 deaths (152 men and 124 women). Overall, the mean age of the patients was 65 years old (mean age 68 ± 15 years for the group who died and 65 ± 17 years for the survivor group). More details of the patients are presented in Table 1.

Studies have shown that APACHE II scoring system alone is not useful for predicting patient mortality, classification of the severity of disease, and duration of stay of the patients⁶, but its use in monitoring therapeutic approaches, comparing the effectiveness of treatments and the decision to change the treatment and the comparison of the functions and quality of the services provided is helpful⁷. It should be noted that some studies have found that this grading system is useful in the classification of patients admitted to ICU and better management of patients who have undergone surgery⁸.

A study used REMS model, which is actually a summarized version of APACHE- II model, to predict the mortality of patients with trauma. Moreover, 3680 trauma patients entered the study over a 4-year period, the result of REMS model was compared with the trauma rating system called RTS. The result of this study showed that REMS model had differentiation equal to RTS, and the area of ROC curve was 0.91 and 0.89, respectively⁹. Another study was conducted on patients with hepatic portal venous gas (HPV) and data from 66 patients with HPVG from EDs 2 were analyzed. REMS, RAPS and MEWS were calculated based on ED data and mortality rates were calculated based on these scores

for each patient. The capability of REMS and RAPS were evaluated to estimate group mortality using ROC curve analysis and calibration analysis. The results showed that sensitivity, specificity and accuracy for each scoring system were 92.1 per cent and 89.3 percent for REMS, 86.8 per cent and 82.1 per cent for RAPS, respectively, 78.9 per cent and 89.3 per cent, respectively. In the analysis of ROC curve, the areas under REMS and RAPS curves were 0.929 and 0.877, respectively. The results showed that REMS predicts mortality in these patients better than RAPS¹⁰. With the superiority of REMS model confirmed in this study as well.

Conclusion

The results showed that APACHE II model has a higher degree of differentiation than other models. Moreover, the overall performance of all systems was in the acceptable range, but the compliance of the predicted actual mortality rate was relatively low. In other words, the estimations showed poor calibration of the model. As these systems have the potential to improve clinical decisions, and as these models are European and American, recalibration of the models seems necessary.

References

1. Olsson T, Terént A, Lind L. Rapid Emergency Medicine Score: a new prognostic tool for in-hospital mortality in nonsurgical emergency department patients. *J Intern Med.* 2004;255(5):579-87. Doi: <https://doi.org/10.1111/j.1365-2796.2004.01321.x>
2. Brabrand M, Folkestad L, Clausen NG, et al. Risk scoring systems for adults admitted to the emergency department: a systematic review. *Scand J Trauma, Resusc Emerg Med.* 2010;18:1-8. Doi: <https://doi.org/10.1186/1757-7241-18-8>
3. Sun BC, Mohanty SA, Weiss R, et al. Effects of hospital closures and hospital characteristics on emergency department ambulance diversion, Los Angeles County, 1998 to 2004. *Ann Emerg Med.* 2006;47(4):309-16. Doi:<https://doi.org/10.1016/j.annemergmed.2005.12.003>
4. Richardson DB. Increase in patient mortality at 10 days associated with emergency department overcrowding. *Med J Aust.* 2006;184(5):213-6. Doi: <https://doi.org/10.5694/j.1326-5377.2006.tb00204.x>
5. Gance LG, Osler T, Shinozaki T. Effect of varying the case mix on the standardized mortality ratio and W statistic: a simulation study. *Chest.* 2000;117(4):1112-7. Doi:<https://www.sciencedirect.com/science/article/pii/S0012369215328154>
6. Zhang Z, Chen K, Chen L. APACHE III outcome prediction in patients admitted to the intensive care unit with sepsis associated acute lung injury. *PLOS ONE.* 2015;10(9):e0139374. Doi: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0139374>
7. Ratanarat R, Thanakittiwirun M, Vilaichone W, et al. Prediction of mortality by using the standard scoring systems in a medical intensive care unit in Thailand. *J Med Assoc Thai.* 2005;88(7):949-55.
8. Khwannimit B, Geater A. A comparison of APACHE II and SAPS II scoring systems in predicting hospital mortality in Thai adult intensive care units. *J Med Assoc Thai.* 2007;90(4):643.
9. Imhoff BF, Thompson NJ, Hastings MA, et al. Rapid Emergency Medicine Score (REMS) in the trauma population: a retrospective study. *BMJ open.* 2014;4(5):e004738. Doi: <http://dx.doi.org/10.1136/bmjopen-2013-004738>.
10. Seak CJ, Yen DH, Ng CJ, et al. Rapid Emergency Medicine Score: A novel prognostic tool for predicting the outcomes of adult patients with hepatic portal venous gas in the emergency department. *PLOS ONE.* 2017;12(9):e0184813. Doi: <https://doi.org/10.1371/journal.pone.0184813>

PEER REVIEW

Not commissioned. Externally peer reviewed.

CONFLICTS OF INTEREST

The authors declare that they have no competing interests.

FUNDING

None

ETHICS COMMITTEE APPROVAL

Ethics committee of Mashhad University of Medical Science approval reference number is IR.MUMS.fm.REC.1395.637

Evaluation of APACHE II, RAPS, and REMS models at the emergency department to predict hospital death in triage levels 1 and 2

Hamed Shokoohsaremi*, Mohsen Ebrahimi, Azade Shirinzadeh Feizabadi

Department of Emergency Medicine, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad

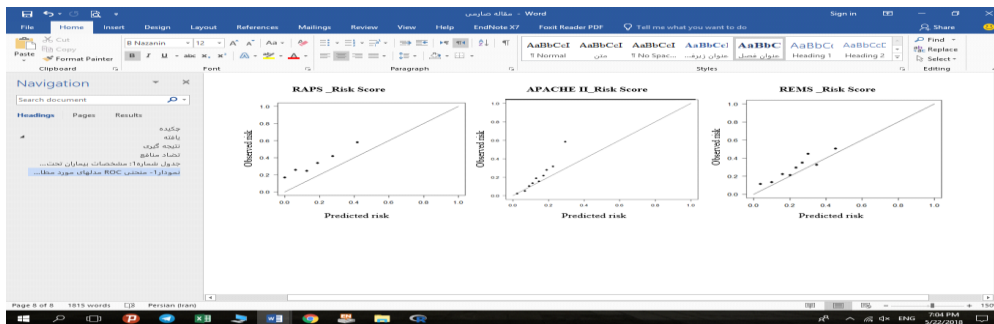


Figure 1: Receiver operating characteristic curves (ROC curves). ROC curves for APACHE II, RAPS and REMS models.

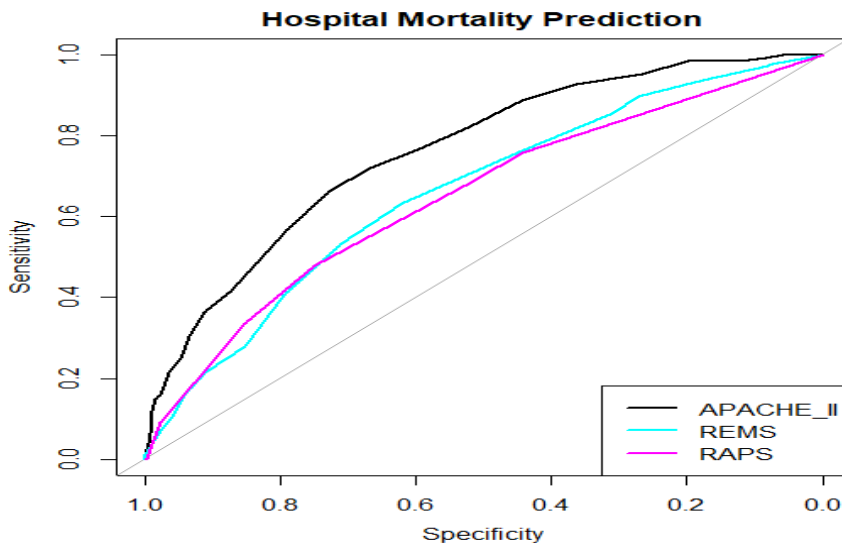


Figure 2: Calibration plots of the APACHE II, RAPS and REMS models in emergency department.

Table 1: Baseline clinical characteristics of patients.

Characteristics		Alive	Dead	P-Value
Gender	Male	152	413	0.501
	Female	124	340	
Triage	Level 1	77	121	<0.001
	Level 2	199	632	
Age		63.1±17.6	68.94±16.01	<0.001
Pulse		94.7±20.4	101.28±22.08	<0.001
MAP		94.6±18.7	89.81±20.97	<0.001
Temp		37.2±0.8	37.23±0.87	0.99
RR		20.4±5.9	22.25±7.17	<0.001
O ₂ Saturation		93.1±5.9	90.68±7.43	<0.001
GCS		14.4±1.3	13.15±2.25	<0.001
HCT		34.9±9.4	34.99±9.62	0.947
WBC		11.0±9.0	14.55±14.88	<0.001
Cr		1.9±2.3	2.48±2.30	0.001
Urine Output		1431.6±281	1296.3±399	<0.001
Na		136.6±6.5	137.03±8.42	0.399
K		4.3±0.9	4.65±1.15	<0.001
Alb		3.6±0.5	3.25±0.65	<0.001
Bs		150.0±94.2	180.69±125	<0.001
Urea		70.8±61.6	114.14±85.2	<0.001
ALT		98.38±54	112±55	<0.001
ALP		152±54.94	180±69.4	<0.001
LDH		256±64.6	394±94.5	0.04
Neut		75.5±15.23	82±15.5	<0.001
Leukocyte		17.20±12.33	12.82±13.7	<0.001

Table 2: Diagnostic value of APACHE II, RAPS and REMS models to predict hospital mortality in emergency department

Risk score system	Brier Score	Area	Std. Error	Confidence Interval		Calibration	
				Lower Bound	Upper Bound	Hosmer Lemshow_Test	P_value
APACHE II	0.141	0.76	0.16	0.728	0.792	113	<0.05
RAPS	0.208	0.629	0.20	0.590	0.669	410	<0.05
REMS	0.186	0.669	0.19	0.632	0.706	46	<0.05

**Evaluation of APACHE II, RAPS, and REMS models at the emergency
department to predict hospital death in triage levels 1 and 2**

Hamed Shokoohsaremi*, Mohsen Ebrahimi, Azade Shirinzadeh Feizabadi

Department of Emergency Medicine, Faculty of Medicine, Mashhad University of Medical Sciences,
Mashhad