Retrospective Study

Scoring systems in critically ill: Which one to use in cancer patients?

Anisha Beniwal, Deven Juneja, Omender Singh, Amit Goel, Akhilesh Singh, Hemant Kumar Beniwal

Abstract

BACKGROUND
Scoring systems have not been evaluated in oncology patients. We aimed to assess the performance of Acute Physiology and Chronic Health Evaluation (APACHE) II, APACHE III, APACHE IV, Simplified Acute Physiology Score (SAPS) II, SAPS III, Mortality Probability Model (MPM) II, and Sequential Organ Failure Assessment (SOFA) score in critically ill oncology patients.

AIM
To compare the efficacy of seven commonly employed scoring systems to predict outcomes of critically ill cancer patients.

METHODS
We conducted a retrospective analysis of 400 consecutive cancer patients admitted in the medical intensive care unit over a two-year period. Primary outcome was hospital mortality and the secondary outcome measure was comparison of various scoring systems in predicting hospital mortality.

RESULTS
In our study, the overall intensive care unit and hospital mortality was 43.5% and 57.8%, respectively. All of the seven tested scores underestimated mortality. The mortality as predicted by MPM II, predicted death rate (PDR) was nearest to the actual mortality followed by that predicted by APACHE II, with a standardized mortality rate (SMR) of 1.305 and 1.547, respectively. The best calibration was shown by the APACHE III score ($\chi^2 = 4.704, P = 0.788$). On the other hand, SOFA score ($\chi^2 = 15.966, P = 0.025$) had the worst calibration, although the difference was not statistically significant. All of the seven scores had acceptable discrimination with good efficacy however, SAPS III PDR and MPM II PDR (AUROC = 0.762),
had a better performance as compared to others. The correlation between the different scoring systems was significant \( (P < 0.001) \).

**CONCLUSION**

All the severity scores were tested under-predicted mortality in the present study. As the difference in efficacy and performance was not statistically significant, the choice of scoring system used may depend on the ease of use and local preferences.

**Key Words:** APACHE score; Intensive care unit; Medical oncology; SOFA score; Scoring systems; Severity of illness index

©The Author(s) 2022. Published by Baishideng Publishing Group Inc. All rights reserved.

---

**Core Tip:** Scoring systems are important for patient triaging, benchmarking intensive care unit (ICU) performance, comparing different ICUs and may also help in patient prognostication, selecting treatment options and resource utilization. However, validity and utility of these scores may be questionable in the patient population apart from where they were developed. Hence, these scores need to be tested and validated in different patient populations, in different geographical areas and over different time periods. There is a lack of an ideal score for prognostication of critically ill cancer patients. In our retrospective study, analyzing data from 400 patients and comparing seven commonly employed critical illness scores, we observed that all the scores had similar efficacy and under-predicted mortality. Therefore, the selection of severity of illness score should depend on the ease of use and local preferences.

**Citation:** Beniwal A, Juneja D, Singh O, Goel A, Singh A, Beniwal HK. Scoring systems in critically ill: Which one to use in cancer patients? *World J Crit Care Med* 2022; 11(6): 364-374

**URL:** [https://www.wjgnet.com/2220-3141/full/v11/i6/364.htm](https://www.wjgnet.com/2220-3141/full/v11/i6/364.htm)

**DOI:** [https://dx.doi.org/10.5492/wjccm.v11.i6.364](https://dx.doi.org/10.5492/wjccm.v11.i6.364)

---

**INTRODUCTION**

The application of prognosticating scoring systems is considered as an important phase in intensive care units (ICUs) since these severity scoring systems estimate the probability of mortality for patients. These scores help the physicians to facilitate resource utilization or continuous quality improvement and to stratify the patients for clinical research[1,2]. ICU scoring systems can help both patients as well as their attendants to select from further treatment options. Further, the scores calculated by these scoring systems help in evaluating the impact of newer treatment modalities and organizational changes which in turn contributes towards the development of treatment standards. In addition to the above, the scoring systems’ outcomes also help in benchmarking ICU performance and comparing the scores secured by different ICU patient populations so as to find out the differences in mortality. However, these systems are unreliable in predicting the clinical outcomes of an individual though it has proven efficacy in predicting mortality for a particular patient cohort[3].

Acute Physiology and Chronic Health Evaluation (APACHE) II and Simplified Acute Physiology Score (SAPS) II are arguably the two most-commonly used and validated tools used in the prediction of ICU patient outcomes[4,5]. These scoring systems were developed in the 1980s and have become outdated due to technological and clinical advancements in critical care management of patients in recent years. Hence, there is a need to develop new scoring systems that include APACHE IV, SAPS III and Mortality Probability Model (MPM) II[6-9]. Such newly-created systems encompass a large number of variables and are highly complicated to compute.

In addition, both validity and utility of the existing scoring systems may be questionable in terms of current patient population compared to the patient population during which they were developed. These scores are widely used and the scoring systems have been validated for a notable time to predict the outcome in general medical or surgical procedures conducted upon critically ill patients. However, whether these systems can predict the mortality accurately among cancer patients remains unknown [10]. There is a dearth of studies that compare different generations of scoring systems and especially the ones used upon cancer patients admitted in medical oncology ICUs. Only a few studies have assessed their usefulness in cancer patients with conflicting results. Moreover, geographic variations in patient populations and the types of cancer necessitate that these scores should be evaluated for different populations[11]. Therefore, the current study is aimed at analyzing the efficacy of seven commonly-used scoring systems to predict the mortality amongst patients admitted in oncology ICUs.
MATERIALS AND METHODS
A retrospective observational cohort study was carried out at a multi-disciplinary onco-medical ICU of a tertiary care center in India. We have an advanced ICU setup and 24-h intensivist coverage with state-of-the-art facilities. Approval for the study and a consent waiver from the institutional ethics committee was obtained.

The data from the records of adult patients who were admitted between January 2018 and February 2020, i.e., 2 years, was collected and analyzed. If the patient was readmitted to the ICU more than once during his/her hospital stay, only the first admission was included in the study. Patients who had ICU stays of less than 12 h, post-operative patients and those admitted from or discharged to another ICU were excluded from the study. Patients fulfilling inclusion criteria were serially recruited. The researchers collected the following data; baseline patient characteristics, indication for ICU admission, type of malignancy, presence of metastasis, need for vasopressor, renal and mechanical ventilation (MV), length of ICU and hospital stay, and ICU and hospital mortality. The data, required to compute various scores, was collected and calculated specified by the procedures.

Statistical analysis
The collected data was then transformed into variables, coded and entered in Microsoft Excel. Then, it was statistically analyzed using SPSS software (version. PC-25). Quantitative data was expressed in mean ± SD or median with an interquartile range. Normality distribution difference between two comparable groups was measured using student’s t-test or Mann Whitney ‘U’ test. Qualitative data was expressed in percentage whereas the statistical differences between the proportions were tested using chi square or Fisher’s exact test, as appropriate.

Standardized Mortality Ratio (SMR) was computed by dividing the observed 28 d’ mortality by predicted hospital mortality based on different scores. Further, 95% confidence interval (CI) was calculated for SMR by considering the observed mortality as a Poisson variable and then dividing its 95% CI by predicted mortality.

The calibration of the scores was executed using Hosmer-Lemeshow goodness-of-fit statistics which divides the subjects into deciles based on the predicted probabilities of death. Afterwards, it computes a Chi-square value from the observed and expected frequencies. Low Chi-square values and high P values (P > 0.5) correspond to a better fit. The ability of the scores to predict ICU mortality was explored and discrimination was tested using Area Under Receiver Operating Characteristic (AUROC) curves. If the AUROC curves are more than 0.8, it denotes excellent outcome while 0.6-0.8 are considered to be acceptable. The cut-off values were calculated for different scores using Youden’s index based on which sensitivity and specificity of the scores were calculated.

Clinically-relevant variables that produced P < 0.05 during univariate analyses and are easily accessible on admission were also entered into multiple logistic regression models as the outcome variable of interest. Odds ratio (OR) was calculated along with 95%CI. A P value < 0.05 was considered to be statistically significant.

Sample size calculation
The sample size calculation was done for the estimation of the AUROC curve for APACHE 2 score, using the following formula:

\[ n \geq \frac{Z_{\alpha/2}^2 \times V(AUC) + d^2}{d^2} \]

Where, \( V(AUC) = 0.0099 \times e^{-a^2/2} \times (6a^2 + 16) \), \( a = \phi^{-1}(AUC) \times 1.414 \) and \( \phi^{-1} \) is the inverse of standard cumulative normal distribution for AUC.

For a 95% level of confidence \( Z_{0.025} = 1.96 \); \( d = 0.05 \) which is the margin of error in estimation and AUC was obtained from a similar study conducted by Schellongowski et al[12] who reported an AUC of 0.776 for the APACHE II score.

Substituting these values in the above formula gives \( n \geq 196 \). As our study was retrospective in nature, we included 400 patients.

RESULTS
During the study period, the data from 400 patients who fulfilled the inclusion criteria were included in the final analysis. Thirty-eight patients were excluded because 31 were admitted from or discharged to another ICU, five were post-operative patients and two had ICU stays less than 12 h. Their baseline characteristics are given in Table 1 and the comparison between various scores is given in Table 2.

Predicted mortality
All of the scoring systems tested in the current study underestimated the mortality (Table 3). The mortality, predicted by MPM II, PDR, was nearest to the actual mortality with an SMR of 1.305, followed by APACHE II (1.547) and SAPS II (1.74).
Table 1 Comparison of baseline variables among survivors and non-survivors

| Parameters                      | Survivors, n = 169 | Non-survivors, n = 231 | Total, n = 400 | P value |
|--------------------------------|--------------------|------------------------|----------------|---------|
| Age in yr                      | 62.85 ± 12.49      | 61.45 ± 14.82          | 62.04 ± 13.88  | 0.527   |
| Male                           | 98 (58.0%)         | 142 (61.5%)            | 240 (60.0%)    | 0.48    |
| Female                         | 71 (42.0%)         | 89 (38.5%)             | 160 (40.0%)    |         |
| DM                             | 56 (33.1%)         | 62 (26.8%)             | 118 (29.5%)    | 0.17    |
| Hypertension                   | 61 (36.1%)         | 63 (27.3%)             | 124 (31.0%)    | 0.06    |
| Reason for ICU admission       |                    |                        |                |         |
| Sepsis                         | 42 (24.9%)         | 68 (29.4%)             | 110 (27.5%)    | 0.31    |
| Respiratory distress/failure   | 76 (45.0%)         | 93 (40.3%)             | 169 (42.2%)    | 0.34    |
| Cardiac arrest                 | 1 (0.6%)           | 8 (3.5%)               | 9 (2.2%)       | 0.08    |
| Gastrointestinal bleed         | 15 (8.9%)          | 14 (6.1%)              | 29 (7.2%)      | 0.33    |
| Altered sensorium              | 33 (19.5%)         | 45 (19.5%)             | 78 (19.5%)     | 1       |
| Acute kidney injury            | 2 (1.2%)           | 3 (1.3%)               | 5 (1.2%)       | 1       |
| Type of malignancy             |                    |                        |                |         |
| Solid organ                    | 135 (79.9%)        | 187 (81.0%)            | 322 (80.5%)    | 0.78    |
| Hematological                  | 34 (20.1%)         | 44 (19.0%)             | 78 (19.5%)     |         |
| Metastasis                     | 80 (59.3%)         | 145 (77.5%)            | 225 (69.9%)    | 0.001   |
| Previous history of surgery for CA |                |                        |                |         |
| Yes                            | 72 (42.6%)         | 74 (32.0%)             | 146 (36.5%)    | 0.03    |
| No                             | 97 (57.4%)         | 157 (68.0%)            | 254 (63.5%)    |         |
| ICU stay                       | 5 (3-8)            | 4 (2-10)               | 5 (3-9)        | 0.58    |
| Hospital stay                  | 14 (8-21)          | 11 (5-22)              | 12 (7-21)      | 0.006   |
| Use of MV                      | 24 (14.2%)         | 130 (56.3%)            | 154 (38.5%)    | < 0.001 |
| Days of MV                     | 5 (3-7.75)         | 3 (2-6)                | 3 (2-7)        | 0.002   |
| Use of renal support           | 7 (4.1%)           | 29 (12.6%)             | 36 (9.0%)      | 0.004   |
| Days of renal support          | 2.14 ± 0.90        | 2.48 ± 2.06            | 2.42 ± 1.88    | 0.786   |
| Use of vasopressor support     | 26 (15.4%)         | 174 (75.3%)            | 200 (50.0%)    | < 0.001 |
| Days of vasopressor support    | 3 (2-4)            | 2 (1.75-4.0)           | 2 (2-4)        | 0.276   |

ICU: Intensive care unit; MV: Mechanical ventilation.

Calibration
Using the Lemeshow-Hosmer goodness-of-fit test, APACHE III (4.704) achieved the best calibration with $P = 0.788$ whereas SOFA score (15.966) was the worst with $P = 0.025$ (Table 4). The least statistically significant discrepancy between the predicted and observed mortality was shown by the APACHE III score.

Discrimination
The efficacy of various scores is given in Figure 1. All the scores tested in the current study exhibited good efficacy, even though there was no statistically significant difference between AUROCs and SAPS III PDR. On the other hand, MPM II PDR (AUROC = 0.762) yielded the best performance (Table 5).

Correlation between various scoring systems
As shown in Table 6, there was a significant correlation found among various scoring systems ($P < 0.001$) as assessed by linear regression analysis.

Factors associated with hospital mortality
Five factors that showed significance in univariate analysis such as hypertension, surgery for cancer, use
Table 2 Comparison between survivors and non-survivors for various scores

| Scoring system     | Survivors, n = 169 | Non-survivors, n = 231 | Total, n = 400 | P value |
|--------------------|--------------------|------------------------|----------------|---------|
| APACHE II          | 17.66 ± 4.96       | 22.82 ± 8.34           | 20.64 ± 7.55   | < 0.001 |
| APACHE II PDR      | 28.10 ± 17.74      | 44.04 ± 25.88          | 37.30 ± 24.10  | < 0.001 |
| APACHE III         | 59.01 ± 16.95      | 81.36 ± 31.37          | 71.92 ± 28.46  | < 0.001 |
| APACHE III PDR     | 17.59 ± 15.80      | 37.59 ± 28.51          | 29.14 ± 25.91  | < 0.001 |
| APACHE IV          | 58.80 ± 16.98      | 80.45 ± 31.70          | 71.30 ± 28.55  | < 0.001 |
| APACHE IV PDR      | 20.45 ± 14.99      | 40.45 ± 27.91          | 32.00 ± 25.33  | < 0.001 |
| SAPS II            | 34.67 ± 11.83      | 49.20 ± 19.87          | 43.06 ± 18.39  | < 0.001 |
| SAPS II PDR        | 19.81 ± 16.97      | 42.83 ± 30.51          | 33.10 ± 28.06  | < 0.001 |
| SAPS III PDR       | 18.12 ± 16.95      | 34.66 ± 24.12          | 27.67 ± 22.88  | < 0.001 |
| SOFA Score         | 5.76 ± 2.80        | 9.02 ± 4.58            | 7.64 ± 4.24    | < 0.001 |
| MPM II PDR         | 33.39 ± 15.08      | 52.16 ± 26.63          | 44.23 ± 24.31  | < 0.001 |

APACHE: Acute Physiology and Chronic Health Evaluation; SAPS: Simplified Acute Physiology Score; SOFA: Sequential Organ Failure Assessment; MPM: Mortality Probability Model; PDR: Predicted death rate.

Table 3 Comparison of the actual and predicted mortality rates for the various scoring systems

| Scoring system     | Actual mortality | Predicted mortality | SMR      | 95%CI       |
|--------------------|------------------|---------------------|----------|------------|
| APACHE II          | 0.577            | 0.373               | 1.547    | 1.425-1.678|
| APACHE III         | 0.577            | 0.291               | 1.982    | 1.824-2.151|
| APACHE IV          | 0.577            | 0.320               | 1.803    | 1.659-1.956|
| SAPS II            | 0.577            | 0.331               | 1.743    | 1.604-1.891|
| SAPS III           | 0.577            | 0.277               | 2.083    | 1.917-2.26 |
| MPM II PDR         | 0.577            | 0.442               | 1.305    | 1.201-1.416|

SMR: Standardized mortality rate; CI: Confidence interval; APACHE: Acute Physiology and Chronic Health Evaluation; SAPS: Simplified Acute Physiology Score; MPM: Mortality Probability Model; PDR: Predicted death rate.

of MV, vasopressors and renal support were used in multivariate analysis as well. Out of the five factors, two factors, i.e. need for MV (OR 2.437, 95%CI = 1.315-4.515, P = 0.005) and vasopressor support (OR 10.465, 95%CI = 5.901-18.557, P = 0.000) were statistically associated with hospital mortality.

**DISCUSSION**

The current study compared various mortality prediction scoring systems and found that all the scores under-predicted the mortality in critically-ill cancer patients. Amongst the scoring systems considered, mortality predicted by MPM PDR was the closest to that of the actual mortality with an SMR of 1.305. AUROC values showed that all of the seven scoring systems had good efficacy and acceptable discrimination. MPM PDR and SAPS III PDR achieved the best discrimination. We found the best sensitivity in SAPS II score (76.2%) and best specificity in SAPS III PDR score (92%). The Lemeshow-Hosmer goodness-of-fit tests showed that the APACHE III score had the best calibration although there was no statistically significant difference.

In the current study, all of the scores were significantly higher among non-survivors (P value < 0.001) as reported in the literature[13-18]. However, all the scores tested in this study underestimated the mortality (SMR > 1), like previous studies[14,15,19,20].

Discrimination is the ability to determine the patients who may die and who will survive. Measures of discrimination include sensitivity, specificity and AUROC curve. But no single scoring system excelled in all of the three areas. SAPS III PDR and MPM II PDR (AUROC = 0.762) had the best AUROC values whereas sensitivity was at its best for SAPS II and specificity was at its best for SAPS III PDR.
Table 4 Lemeshow-Hosmer goodness-of-fit tests for evaluating the calibration of the scoring systems

| Scoring system     | Chi square value | P value |
|--------------------|------------------|---------|
| APACHE II          | 9.366            | 0.312   |
| APACHE II PDR      | 12.159           | 0.144   |
| APACHE III         | 4.707            | 0.788   |
| APACHE III PDR     | 6.471            | 0.595   |
| APACHE IV          | 9.331            | 0.315   |
| APACHE IV PDR      | 10.763           | 0.216   |
| SAPS II            | 9.479            | 0.304   |
| SAPS II PDR        | 10.410           | 0.237   |
| SAPS III PDR       | 10.787           | 0.214   |
| SOFA Score         | 15.966           | 0.025   |
| MPM II, PDR        | 11.265           | 0.187   |

APACHE: Acute Physiology and Chronic Health Evaluation; SAPS: Simplified Acute Physiology Score; SOFA: Sequential Organ Failure Assessment; MPM: Mortality Probability Model; PDR: Predicted death rate.

Table 5 Area under curve for predicting hospital mortality for various scoring system

| Scoring system     | AUC   | P value | 95%CI     | Cut off | Sensitivity | Specificity |
|--------------------|-------|---------|-----------|---------|-------------|-------------|
| APACHE II          | 0.688 | < 0.001 | 0.637-0.739 | > 18.5  | 67.5%       | 62.7%       |
| APACHE III         | 0.720 | < 0.001 | 0.672-0.769 | > 78.5  | 46.8%       | 87.6%       |
| APACHE IV          | 0.708 | < 0.001 | 0.659-0.758 | > 72.5  | 53.7%       | 79.3%       |
| SAPS II            | 0.734 | < 0.001 | 0.685-0.782 | > 34.5  | 76.2%       | 60.4%       |
| SAPS III PDR       | 0.762 | < 0.001 | 0.715-0.808 | 39.0    | 44.3%       | 92.0%       |
| SOFA Score         | 0.715 | < 0.001 | 0.665-0.764 | > 7.5   | 58.0%       | 79.3%       |
| MPM II, PDR        | 0.762 | < 0.001 | 0.714-0.810 | 36.45   | 71.3%       | 69.9%       |

AUC: Area under the curve; CI: Confidence interval; APACHE: Acute Physiology and Chronic Health Evaluation; SAPS: Simplified Acute Physiology Score; SOFA: Sequential Organ Failure Assessment; MPM: Mortality Probability Model; PDR: Predicted death rate.

However, these differences were not statistically significant. In the current study, AUROC outcomes showed that discrimination is acceptable in all the scoring systems tested as reported in the literature [14-16,20-22]. All the severity illness scores showed good efficacy with no statistically significant difference in AUROCs.

Calibration evaluates the accuracy of the degree of correspondence between the estimated probability of mortality and the observed actual mortality. Calibration is good if the predicted mortality is close to the observed mortality. APACHE III (4.704) had the best calibration with P = 0.788. This infers that it had the least statistically significant discrepancy between the predicted and observed mortality. Good calibration of these scores have also been reported by other authors [14-16,20].

A significant correlation was found among various scoring systems (P < 0.001) as per linear regression analysis. This correlation may be attributed to the overlap of multiple variables, considered for calculating the scores. Sculier et al.[21] also reported an excellent correlation between APACHE II and SAPS II in their study on oncology patients. ICU mortality rate among cancer patients was reportedly high and in the range of 30% to 77% [23-26]. The overall ICU mortality rate in the current study was 43.5%. Even though it is higher, the ICU mortality of the current cohort does not differ from the mortality reported in similar studies conducted earlier [23,24]. The hospital mortality rate in the current study was 57.8% which is again similar as reported earlier [27,28].

Use of MV and vasopressor support have a direct association with hospital mortality. Similar studies conducted earlier have also reported the need for organ support in the form of MV. At times, vasopressor use is directly associated with increased mortality among cancer patients [29]. An ideal scoring system is the need of the hour. This system should be well calibrated, easy to compute, able to have high
levels of discrimination and predict mortality rates with high accuracy based on the easily-available patient parameters. Additionally, an ideal score also needs to be dynamic, reflecting the change in management and case mix over time. In this search for an ideal scoring system, newer scoring systems have been developed. However, these systems are highly complex in nature, demand huge sets of patient data and need computer assistance to calculate the scores. Hence, the development of an ideal scoring system has a long way to go.

The accuracy of scoring systems may differ over a period of time and may produce varied results in different countries due to differences in ethnicity, patient population, healthcare systems, ICU structure and organization. So, its accuracy cannot be generalized and all such models need external validation in independent patient populations to prove its reproducibility. Therefore, it becomes imperative to compare and test the validity of scoring systems under different geographical areas and upon different patient populations. The current study is one of the few studies conducted on the Indian subcontinent and the researchers have compared a huge number of scoring systems developed for cancer patients in a significantly large cohort of patients.
The current study has a limitation to address, i.e. being a single center retrospective study where concerns may arise in terms of generalizing the conclusions arrived in this study. The missing data may have also led to information bias. Nonetheless, the study has several salient features such as the comparison of seven scoring systems, fairly large sample size, well-defined study protocol and the inclusion of only medical oncology patients.

CONCLUSION

The current study concludes that all of the scoring systems considered for this study cohort under-predicted the mortality. However, the APACHE III score had the least discrepancy between the predicted and observed mortality. There was no statistically significant difference in efficacy and all the scores tested had good calibration and acceptable discrimination. Hence, the choice of scoring system in critically-ill oncology patients should not only be based on the performance of the score, but also on other factors such as ease of use and local preferences.

ARTICLE HIGHLIGHTS

Research background

The application of prognosticating scoring systems is considered as an important phase in intensive care units (ICUs) since these severity scoring systems estimate the probability of mortality for patients. These scores help the physicians to facilitate resource utilization or continuous quality improvement and to stratify the patients for clinical research. ICU scoring systems can help both patients as well as their attendants to select from further treatment options. Further, the scores calculated by these scoring systems help in evaluating the impact of newer treatment modalities and organizational changes which in turn contributes towards the development of treatment standards. In addition to the above, the scoring systems’ outcomes also help in benchmarking ICU performance and comparing the scores secured by different ICU patient populations so as to find out the differences in mortality.

Research motivation

There is a dearth of studies that compare different generations of scoring systems especially the ones used upon cancer patients admitted in medical oncology ICUs. Only a few studies have assessed their usefulness in cancer patients with conflicting results.

Research objectives

To compare the efficacy of seven commonly employed scoring systems to predict outcomes of critically ill cancer patients.
Research methods
We conducted a retrospective analysis of 400 consecutive cancer patients admitted in the medical intensive care unit over a 2-year period. The primary outcome was hospital mortality and the secondary outcome measure was comparison of various scoring systems in predicting hospital mortality.

Research results
Overall ICU mortality in our study was 43.5% whereas hospital mortality was 57.8%. All scoring systems tested underestimated the mortality. Mortality predicted by MPM II, predicted death rate (PDR), was closest to that of the actual mortality followed by that of APACHE II, with a standardized mortality rate (SMR) of 1.305 and 1.547, respectively. APACHE III ($\chi^2 = 4.704, P = 0.788$) had the best calibration and SOFA score ($\chi^2 = 15.966, P = 0.025$) had the worst calibration, but the difference was not statistically significant. All the scores tested had good efficacy and acceptable discrimination, however SAPS III PDR and MPM II PDR (AUROC = 0.762), performed better than others. There was a significant correlation between the various scoring systems ($P < 0.001$).

Research conclusions
Overall, all the scores in our study cohort under-predicted the mortality. The difference in efficacy was not statistically significant in all scores. The choice of scoring system should depend on the ease of use and local preferences as all the scores tested had similar performance.

Research perspectives
There is a lack of an ideal score for prognostication of critically ill cancer patients. In our retrospective study, analyzing data from 400 patients and comparing seven commonly employed critical illness scores, we observed that all the scores had similar efficacy but under-predicted mortality. Therefore, the choice of scoring system should depend on the ease of use and local preferences.

FOOTNOTES

Author contributions: Beniwal A and Juneja D designed the study; Beniwal A, Juneja D and Beniwal HK collected the data, analyzed the results, performed the majority of the writing and prepared the tables; Singh O, Goel A and Singh A provided critical input in writing the paper and reviewed the manuscript.

Institutional review board statement: Approved by Institutional Scientific Committee of Max Super Speciality Hospital, No. 1944105991.

Informed consent statement: As this was a retrospective study, the need for consent was waived off by the institute’s ethical committee.

Conflict-of-interest statement: All authors report no relevant conflict of interest for this article.

Data sharing statement: No additional data are available.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: https://creativecommons.org/Licenses/by-nc/4.0/

Country/Territory of origin: India

ORCID number: Anisha Beniwal 0000-0003-1139-0692; Deven Juneja 0000-0002-8841-5678; Omender Singh 0000-0002-3847-4645; Amit Goel 0000-0002-9509-5705; Akhilesh Singh 0000-0001-6192-140X; Hemant Kumar Beniwal 0000-0002-7052-0146.

S-Editor: Wu YXJ
L-Editor: Filipodia
P-Editor: Wu YXJ

REFERENCES
1 Sakr Y, Krauss C, Amaral AC, Réa-Neto A, Specht M, Reinhart K, Marx G. Comparison of the performance of SAPS II, SAPS 3, APACHE II, and their customized prognostic models in a surgical intensive care unit. Br J Anaesth 2008; 101: 798-803 [PMID: 18845649 DOI: 10.1093/bja/aen291]
Yoo H, Jeong BH, Lim SY, Chung MP, Kwon OJ, Jeon K. Etiologies, diagnostic strategies, and outcomes of...
diffuse pulmonary infiltrates causing acute respiratory failure in cancer patients: a retrospective observational study. *Crit Care* 2013; 17: R150 [PMID: 23880212 DOI: 10.1186/cc12829]

27 **Wang YG**, Zhou JC, Wu KS. High 28-day mortality in critically ill patients with sepsis and concomitant active cancer. *J Int Med Res* 2018; 46: 5030-5039 [PMID: 30088429 DOI: 10.1177/0300060518789040]

28 **Shin SH**, Lee H, Kang HK, Park JH. Twenty-eight-day mortality in lung cancer patients with metastasis who initiated mechanical ventilation in the emergency department. *Sci Rep* 2019; 9: 4941 [PMID: 30894559 DOI: 10.1038/s41598-019-39671-8]

29 **Kostakou E**, Rovina N, Kyriakopoulou M, Koulouris NG, Koutsoukou A. Critically ill cancer patient in intensive care unit: issues that arise. *J Crit Care* 2014; 29: 817-822 [PMID: 24857640 DOI: 10.1016/j.jcrc.2014.04.007]
