Predicting Burn Mortality Using a Simple Novel Prediction Model

Sneha Sharma¹  Raman Tandon¹

¹Department of Burns, Plastic and Maxillofacial Surgery, Vardhman Mahavir Medical College and Safdarjung Hospital, New Delhi, India

Address for correspondence Raman Tandon, MCh, Department of Burns, Plastic and Maxillofacial Surgery, Vardhman Mahavir Medical College and Safdarjung Hospital, Ansari Nagar, New Delhi 110029, India (e-mail: tandonraman410@gmail.com).

Introduction

Burn-inflicted morbidity and mortality plague societies world over. Thermal injuries are a major cause of public health burden in the developing countries.¹ Most patients sustaining <30% total burn surface area (TBSA), have an uneventful recovery, while patients with extensive burns (>60% TBSA) have a dismal prognosis. Patients with 30 to 60% TBSA have reported mortality of 40%² Most of these patients require intensive care. Accurately predicting an outcome allows the health care providers to appropriately allocate resources, prognosticate family members, and to perform audits and scientific studies.

Abstract

Background  Prediction of outcome for burn patients allows appropriate allocation of resources and prognostication. There is a paucity of simple to use burn-specific mortality prediction models which consider both endogenous and exogenous factors. Our objective was to create such a model.

Methods  A prospective observational study was performed on consecutive eligible consenting burns patients. Demographic data, total burn surface area (TBSA), results of complete blood count, kidney function test, and arterial blood gas analysis were collected. The quantitative variables were compared using the unpaired student t-test/nonparametric Mann Whitney U-test. Qualitative variables were compared using the χ²-test/Fischer exact test. Binary logistic regression analysis was done and a logit score was derived and simplified. The discrimination of these models was tested using the receiver operating characteristic curve; calibration was checked using the Hosmer–Lemeshow goodness of fit statistic, and the probability of death calculated. Validation was done using the bootstrapping technique in 5,000 samples. A p-value of <0.05 was considered significant.

Results  On univariate analysis TBSA (p <0.001) and Acute Physiology and Chronic Health Evaluation II (APACHE II) score (p = 0.004) were found to be independent predictors of mortality. TBSA (odds ratio [OR] 1.094, 95% confidence interval [CI] 1.037–1.155, p = 0.001) and APACHE II (OR 1.166, 95% CI 1.034–1.313, p = 0.012) retained significance on binary logistic regression analysis. The prediction model devised performed well (area under the receiver operating characteristic 0.778, 95% CI 0.681–0.875).

Conclusion  The prediction of mortality can be done accurately at the bedside using TBSA and APACHE II score.
The prediction of burn mortality has come a long way since Bull and Squire first studied the effects of age and TBSA on mortality. Baux et al in 1961 devised a score that added the age and TBSA. These models provided a simplified estimate of mortality, but the observed mortality trends were more nuanced. More meticulous models for the estimation of the probability of death were required and thus, many newer models were developed using complex statistical methods. In the quest for more refined predictions, complexities increased.

We designed this study intending to devise a simple burn-specific mortality prediction model incorporating both exogenous and endogenous factors.

**Materials and Methods**

**Study Design**

This prospective observational study was designed to identify the patient variables, within 24 hours of admission, which predict patient mortality and to formulate a simple to use mortality prediction model.

This study was started after ethical clearance from the Institute Ethics Committee. Patient recruitment was performed over 15 months from December 2018 to February 2020.

All consecutive patients admitted were assessed for eligibility and consenting adult patients, presenting to the burn casualty within 24 hours of sustaining 30 to 60% scald and thermal burns, in the age group of 18 to 55 years were enrolled. Electrical, chemical, radiation, and friction burns were excluded. Burn patients with trauma, comorbidities, and pregnancy were excluded.

Objective estimation of TBSA was made in the burn casualty using a burn chart formulated by the institute following the Lund and Browder chart. The estimation of the depth of the burn was subjective, based on visual examination. The diagnosis of inhalational injury was made using the following history and clinical criteria: (1) history of being involved in a closed space fire, (2) the presence of facial burns with singed nasal and facial hair, (3) production of carbonaceous sputum, and (4) signs of respiratory distress. Routine hematological work-up comprising of complete blood count, kidney function test, and arterial blood gas analysis was sent.

All consecutive patients were approached and assessed for eligibility and consent during the first 24 hours of their admission. After obtaining consent, patient demographics, vitals, and results of the hematological work-up were collected. Using these variables Fatality by Longevity, APACHE II score, Measured Extent of burn and Sex (FLAMES) score was calculated, and the probability of death computed.

The patients were followed at 10, 20, and 30 days for assessing mortality.

**Statistical Analysis**

The data has been presented in terms of frequency and percentage for qualitative variables and mean ± SD for normally distributed data and median (interquartile range) for skewed data and range for quantitative variables.

The normality of the variables was checked using the Kolmogorov-Smirnov test. The quantitative variables between the survivor and nonsurvivor groups were compared using the unpaired student t-test for normally distributed data and nonparametric Mann Whitney U test for skewed data. Qualitative variables were compared using the χ2-test for normally distributed data and Fischer exact test for skewed data. A two-tailed p-value of <0.05 was considered significant.

The variables which were found to be significant during this preliminary analysis were used to formulate a model (to find the probability of death) using binary logistic regression. The general formula for a binary logistic regression equation is:

\[
\text{Logit} = b_0 + b_1x_1 + b_2x_2 + \ldots + b_{nx_n},
\]

Using this logit score the probability of death was calculated using the following equation:

\[
\text{Probability of death} = \frac{e^{\text{logit}}}{1 + e^{\text{logit}}},
\]

The model derived above was further simplified and named the APACHE II–TBSA (AT) score (explained further in the Results section).

The discrimination of these two models (the logit score and the AT score) and the FLAMES score was tested using the receiver operating characteristic (ROC) curve. A p-value of ≤0.05 was considered significant. The ROC was plotted for 10-day, 20-day, and 30-day mortality. Calibration for the three models was tested using the Hosmer–Lemeshow goodness of fit C statistic. The probability of death was calculated using each of the three models. A case with a calculated probability of death ≥0.5 was considered more likely to die and <0.5 was considered more likely to survive. Contour plots for the two models were created and examined.

Validation of the models was done using bootstrapping statistic (sampling and replacement method) in 5,000 samples. A p-value of ≤0.05 was considered significant.

The AUC for the models and FLAMES score were plotted for 10-day, 20-day, and 30-day mortality and compared. A p-value ≤0.05 was considered significant.

Statistical packages SPSS 23 (IBM Corp. Released 2015. IBM SPSS Statistics for Windows, Version 23.0. Armonk, New York, United States: IBM Corp.) and Stata 16 (StatCorp. 2019. Stata Statistical Software: Release 16. College Station, Texas, United States: StataCorp LLC.) were used.

**Results**

From December 2018 to February 2020, 1,940 burns patients were admitted within 24 hours of sustaining thermal injuries. Out of these 1,831 patients were excluded and 109 were enrolled as shown in (Supplementary Fig. S1 [available online only]). Nine patients were lost to follow-up and excluded from the study.

The median age of the patients included was 30 years (25, 40), ranging from 18 to 55 years. Fifty-one were females and 49 males. The median TBSA was 40% (35, 50), ranging from 30 to 60%. Ninety-seven patients sustained thermal burns, while three patients sustained scald burns. A diagnosis
of inhalational injury was made in 84 of the 100 patients. Observed 30-day mortality was 36%.

The characteristics and comparison of the survivor and nonsurvivor groups are as shown in Table 1.

On univariate analysis, Acute Physiology And Chronic Health Evaluation II (APACHE II) score ($p = 0.004$) and TBSA ($p < 0.001$) were found to be significant predictors of mortality (Table 1). Other variables which were found to be significant influencers of mortality were heart rate, alveolar-arterial oxygen gradient ($A-aPaO_2 = \left[\left(\frac{P_{atm} - P_{H_2O}}{\text{FiO}_2}\right) - \left(\frac{P_{CO_2}}{\text{RQ}}\right)\right] - PaO_2$, where $P_{atm}$ is the atmospheric pressure, $P_{H_2O}$ is the partial pressure of water in the alveoli, $\text{FiO}_2$ is the fraction of inspired oxygen, $P_{CO_2}$ is the partial pressure of $CO_2$ in the alveoli, $\text{RQ}$ is the respiratory quotient and $PaO_2$ is the arterial oxygen pressure, serum creatinine levels, and Glasgow coma scale ($p$-values $<0.05$). These variables were not included for further analysis as these are already considered in the APACHE II score.

Binary logistic regression was performed and a model using the APACHE II score and TBSA was formulated (Table 2). This yielded an equation that was used to derive a logit score, based on which probability of death could be calculated.

$$\text{logit} = -5.692 + 0.153\cdot(\text{APACHE II Score}) + 0.090\cdot(\text{TBSA})$$

The formula used to calculate the probability of death from this score is:

$$\text{Probability of death} = \frac{e^{\text{logit}}}{1+e^{\text{logit}}}.$$
This logit score provides a fair estimation of the probability of death with an area under the ROC curve (AUC) of 0.778, p-value <0.001, 95% confidence interval (CI) (0.681, 0.875) (►Fig. 1).

It is intuitive to think that a high TBSA will result in a higher APACHE II score for a particular patient. Our results prove that TBSA and APACHE II score are independent predictors of mortality in burn patients. The variance inflation factor for the model is 1.074 and it implies that there are no significant multicollinearity effects in this model.

Since the calculation of this score is difficult at the bedside, it was simplified. As shown in ►Table 2, the coefficient for the APACHE II score is 1.7 times the coefficient for TBSA. Using 1.7 as the multiplier for the APACHE II score, a simplified model was formulated. This score will be referred to as the AT score (APACHE II–TBSA score):

\[ AT = 1.7(\text{APACHE II score}) + \text{TBSA} \]

Theoretically, AT score can range from zero to 220.7. In this study sample, the scores ranged from 30 to 111.20. This score can be used to provide a rough estimate of predicted mortality of a patient at the bedside, as a higher score is associated with increased mortality. A graph of observed mortality against the score was plotted as shown in ►Fig. 2. There is a nonlinear relationship between the score and the observed mortality. It can be appreciated that a score of ≤50 has a mortality rate of <20%, for a score of 50.1 to 70 the mortality is roughly 40%, a score of 70.1 to 90 has a mortality of 75%, and at a score above 90 the observed mortality reaches 100%.

The performance of the score in predicting mortality was fair with an AUC identical to that of the logit score (0.778, p-value <0.001, 95% CI [0.681, 0.875]) (►Fig. 1).

Even though the discriminative ability of the score is good, it could only provide a rough estimate of the probability of death. For a more accurate prediction of mortality, this model was fitted using binary logistic regression with results as shown in ►Table 3. Using this model, the probability of death can be calculated using the following steps:

1. Calculate “X” using the following logistic regression equation:

\[ X = -5.692 + (0.90 \times \text{AT score}) \]

2. The probability of death can then be calculated using the following formula:

\[ \text{Probability of death} = \frac{e^X}{1 + e^X} \]

In these formulae, e is the Euler’s number with a value of 2.71828.

The performance and calibration of the logit score, the AT score, and the FLAMES score for predicting 30-day mortality are as shown in ►Table 3.

On bootstrapping the models in 5,000 samples, the relationship of APACHE II and TBSA with the 30-day mortality remained significant with a p-value of <0.001.

The predictions of the logit score and AT score are comparable as shown in the contour plots for both in ►Fig. 3.

The performance of the logit score and AT score was compared with that of the FLAMES score in the sample population. The performance for prediction of 10-day, 20-day, and 30-day mortality was compared. The performance of the logit score and AT scores in predicting 10-day and 20-day mortality was significantly better as compared with the performance of the FLAMES score (p-values of 0.0048 and 0.0066, respectively). While for 30-day mortality, the p-values for the AUC are not statistically significant. This implies that the models derived from the sample population have better predictive abilities as compared with the FLAMES score for 10-day and 20-day mortality, while the predictive ability of the three

Table 3  Fitted APACHE II–TBSA score (AT score)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT score</td>
<td>0.090</td>
<td>0.22</td>
</tr>
<tr>
<td>Constant</td>
<td>-5.692</td>
<td>1.271</td>
</tr>
</tbody>
</table>

Abbreviations: APACHE II, Acute Physiology And Chronic Health Evaluation II; TBSA, total burn surface area.
Note: p-Value for all the coefficients is <0.001.

Table 4  Comparison of the performance and calibration of the logit score, APACHE II–TBSA score (AT score) and Fatality by Longevity, APACHE II score, Measured Extent of burn, and Sex (FLAMES) score

<table>
<thead>
<tr>
<th>Model</th>
<th>Logit score</th>
<th>AT score</th>
<th>FLAMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted deaths</td>
<td>29</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td>Hosmer–Lemeshow C-statistic (p-value)</td>
<td>4.915 (0.767)</td>
<td>4.915 (0.767)</td>
<td>9.148 (0.330)</td>
</tr>
<tr>
<td>AUROC (95% CI)*</td>
<td>0.778 (0.681–0.875)</td>
<td>0.778 (0.681–0.875)</td>
<td>0.737 (0.636–0.837)</td>
</tr>
</tbody>
</table>

Abbreviations: APACHE II, Acute Physiology And Chronic Health Evaluation II; AUROC, area under receiver operating characteristic; CI, confidence interval; TBSA, total burn surface area.

*p-Value < 0.001.

Fig. 2 Relationship between values of the APACHE II–TBSA score (AT score) and the observed mortality. APACHE II, Acute Physiology and Chronic Health Evaluation II; TBSA, total burn surface area.

Fig. 3 Relationship between values of the APACHE II–TBSA score (AT score) and the observed mortality. APACHE II, Acute Physiology and Chronic Health Evaluation II; TBSA, total burn surface area.
models is the same when it comes to the 30-day mortality (Table 5 and Fig. 4).

**Discussion**

Our analysis proves that TBSA and APACHE II score are independent predictors of mortality (Table 1). TBSA is a strong predictor of mortality with a p-value <0.001. TBSA is an objective estimate of the severity of thermal injury and has been extensively studied as a predictor of burn-related mortality. TBSA has been included in mortality prediction scores since Weidenfeld first designed the burn-specific mortality score in 1902 and thereafter. Some recent studies concluded that full-thickness burns have a stronger relationship with mortality than the partial thickness burns. Other studies did not find a correlation of burn depth with mortality. Germann et al. found that TBSA was a more important prognostic factor as compared with full-thickness burn area.

In this study, we did not find the depth of burns to be a predictor of mortality. The estimates of burn depth at admission are, at best subjective and prone to interpersonal variability. Burn depth also evolves during hospital stay. The inclusion of objective estimates for variables improves the reliability of the model. Methods have been devised to objectively assess the burn depth as the clinical examination has an accuracy of only 64 to 76% with experienced clinicians. Numerous techniques using vital dyes, nuclear imaging, and thermography have been introduced. The laser doppler technology has proven to be the most viable of these innovations and uses the Doppler principle to assess the blood supply of the burnt tissue and estimate the depth of the burn. Burn depth index is a useful measure helpful in classifying patients into various groups based on the proportion of deep burns that they have and help in predicting the outcome and hence personalizing care.

The other variable found to significantly affect mortality in our study is the APACHE II score. The APACHE II score was...
devised by Knaus et al\textsuperscript{23} as a system for the classification of severity of disease in critically ill patients. This score provides an objective estimate of the host physiological response to injury. The creators of this score did not include thermally injured patients in the study. Through the years, APACHE score has also been studied as a predictor of mortality and disease severity in the burns population.\textsuperscript{8,12,13,24} APACHE II score does not take into account the serum protein and albumin/globulin ratio. These measures have also been proven to be independent mortality predictors in burn patients.\textsuperscript{25}

Using these variables, the logit score was formulated. This score takes into account both the exogenous (measure of injury severity, i.e., TBSA) and endogenous factors (age, heart rate, blood pressure, etc. combined as the APACHE II score) affecting burn-specific mortality. To simplify its use, the logit score was modified to formulate the AT score. It is of note that the simpler model is capable of predicting nuanced patterns of mortality as shown in \textbf{Fig. 3}. The contours of both the plots are similar and show that the predicted mortality has a linear relationship for lower values of the variables and then as these reach a higher value the relationship becomes nonlinear. The AT score retains the ability of the logit score to predict mortality in a clinically more plausible way.

For validation of the model bootstrapping was used. In this statistical test, the ability of a predictive model to give accurate results on data not included in the development sample is checked. The AT score retained the significance of its relationship with mortality suggesting reproducibility.

The FLAMES score\textsuperscript{8} is an already established and validated score.\textsuperscript{13,26} It uses the APACHE II score, percentages of full-thickness burn area, and partial thickness burn area, and sex of the patient for prediction of mortality. It uses a complex logistic regression equation to calculate the score and then predict mortality. On comparison of the FLAMES score with the new models, the AT score was found to be easier to apply at the bedside. The performance of the logit score and the AT score was found to be better for prediction of 10-day and 20-day mortality and similar for 30-day mortality. The Hosmer–Lemeshow C-statistic for the FLAMES score was greater and the \textit{p}-value smaller than that of these scores, implying that even though the FLAMES score is a good fit for the sample population, the newly derived models fare better. The FLAMES score also underestimated the number of deaths in the population.

Despite being a prospective study, this study has its limitations. This study was a single-center study, with stringent inclusion criteria. We included only patients in the young adult and middle-aged categories to reduce the effect of co-morbidities which develop with age. Similarly, patients with preexisting comorbidities, coexistent trauma, and pregnancy were excluded. These conditions have an independent bearing on the mortality of any patient. A thermally injured patient is usually well until the inciting event. To be able to predict mortality accurately in this group of patients the model should be free of these confounding factors at the time of formulation. We only included patients with 30 to 60% TBSA as the prognosis is most uncertain in this group. The abovementioned factors subject the study to sampling bias. Secondly, we did not study the effect of co-morbid illnesses on mortality in the thermally injured patient. But as mentioned earlier, majority of these patients are in good health until the event, so it is possible that patients with comorbidities form a small proportion of this population. Further, sometimes, the co-morbidities are such that they might lead to death in a patient with minor nonfatal injuries. Thirdly, we did not record the complications such as pneumonia, sepsis, etc. But it is generally accepted that these complications have a direct relationship to the severity of the injury and therefore can be seen as a function of TBSA, which has been included in the model. Moreover, the model has been developed to predict the propensity of a patient to die at the time of admission and complications occur later during admission. Fourthly, we did not find the depth of burns to be an independent predictor of mortality in our study and have not taken it into consideration. Lastly, it should be noted that the models fare better in the derivation population, this could be the cause of the better performance of these models as compared with the FLAMES score.

We present these models as easy to use tools that can be used to aid clinical decision making, provide an objective criterion for burn unit audit, and assist further scientific research.

**Conclusion**

TBSA and APACHE II score are important independent predictors of burn-related mortality in American Society of Anesthesiologists Class 1 patients in the age group of 18 to 55 years with 30 to 60% thermal and scald burns. The AT score is an easy to use prognostic model (without the inclusion of burn depth as a variable), with good discriminative ability and calibration, which provides accurate and nuanced predictions of burn-related mortality.

**Declarations**

- Ethics approval and consent to participate: ethics committee approval obtained from the Institutional Ethics Committee.

**References**

Predicting Burn Mortality using a Simple Novel Prediction Model

Sharma, Tandon.


9 McGwin Jr, George RL, Cross JM, Rue LW. Improving the ability to predict mortality among burn patients. Burns 2008;34(3):320–327
22 Kumar P. Burns depth index and classification of burn casualties. Burns 1993;19(3):252
24 Tanaka Y, Shimizu M, Hirabayashi H. Acute physiology, age, and chronic health evaluation (APACHE) III score is an alternative efficient predictor of mortality in burn patients. Burns 2007;33(3):316–320