

Predicting survival of patients with deep burns using supervised machine learning algorithms

Sara Hashemi · Abdolkhalegh Keshavarzi · Leila Erfannia · Abdolrahman Zarei · Azita yazdani 

Received: 6 November 2025 / Accepted: 31 August 2026

© The Author(s) 2026

Abstract

Deep burns carry a high risk of death and early reliable mortality prediction can support triage, resource allocation, and clinical decision-making. We aimed to develop and evaluate a leakage-resistant Machine Learning (ML) model to predict in-hospital mortality in patients with deep burns using only information available within the first 24 h of admission. In this retrospective, single-center study of 528 patients with deep (third-degree) burns (240 in-hospital deaths, 45.5%), twenty-four first-24-hour predictors were analyzed. Ten supervised ML algorithms were compared using nested 10×5 cross-validation, with all preprocessing performed within each fold to prevent leakage. A final untouched 20% hold-out set, derived from the same center and period, was used for evaluation. Class imbalance was addressed with balanced class weights, and missingness not at random was modeled explicitly. All models showed high discrimination. We ultimately selected a regularized logistic regression model, as it combined the highest cross-validated discrimination (area under the receiver-operating-characteristic curve (AUROC) 0.982 ± 0.020) with excellent calibration and full interpretability. On the hold-out set, the model remained robust, achieving an AUROC of 0.992 (95% CI: 0.980–0.999) and a Brier score of 0.04. Notably, while only five of 106 patients were misclassified, the single false negative (FN) highlights a critical need for clinical oversight to prevent missed high-risk cases. Compared to traditional scoring, our model outperformed the Baux (AUROC = 0.943; $p = 0.009$) and revised Baux (AUROC = 0.950; $p = 0.011$) scores, while performing on par with the Abbreviated Burn Severity Index (ABSI) (AUROC = 0.982; $p = 0.22$). Even after Bonferroni correction, the improvements over Baux scores remained statistically significant. The strongest predictor was clearly the percentage of total body surface area burned (%TBSA), followed by early organ-stress markers, inhalation injury, and diabetes. Performance was statistically significant in a permutation test and showed no evidence of bias by sex. The proposed model demonstrates high predictive performance for early mortality among deep-burn patients. Despite being an internal validation study in which burn extent was recorded in ordinal 10% bands rather than exact %TBSA, a limitation that constrains direct clinical translation, the model offers a promising framework for supporting early clinical triage and stratification, provided it undergoes rigorous external multi-center evaluation.

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.

Sara Hashemi and Abdolkhalegh Keshavarzi contributed equally to this work.



early risk stratification, resource allocation, and communication with patients and families, complementing rather than replacing clinical judgement and established scores. These findings should, however, be interpreted as evidence of promise rather than of readiness: the hold-out set was drawn from the same centre and period as the development data and therefore constitutes internal, not external, validation. In addition, burn extent was recorded only as ordinal 10% bands rather than exact %TBSA; this measurement coarseness limits the direct clinical translation of the model and its comparability with cohorts in which exact %TBSA is recorded, and it should be resolved before the model is used in settings that document burn size more precisely. The immediate priority is therefore external validation, and we invite collaboration with other burn centres to that end. Three study designs would be most informative: geographical validation across independent Iranian burn centres to test transportability across case-mix; temporal validation on a later cohort from the same centre to test stability over time; and a prospective study recording exact %TBSA, which would both remove the banding limitation and permit a fair head-to-head comparison with the classical scores. Subsequent work should evaluate the model's effect on clinical decision-making and patient outcomes, and extend prediction to additional endpoints such as length of stay and functional recovery.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-026-70066-8>.

Acknowledgements We thank the Research Deputy of Shiraz University of Medical Sciences for financially supporting this project. We also would like to express our sincere gratitude to all experts who participated in this study. In addition, we gratefully acknowledge the Clinical Research Development Center of Nemazee Hospital for their support in data analysis, and we especially thank Dr. Ali Mohammad Keshtvarz Hesam Abadi for his valuable contribution to the statistical analysis of this work.

Author contributions AY and LE planned the study design and coordinated the conduct of the study. AZ and SA collected the study data and participated in the drafting of the manuscript. SA, AY, and LE carried out the analysis and interpretation of the data and drafted this paper. AY, and AK provided technical counseling on study design and critically revised the manuscript. All authors read and approved of the final manuscript.

Funding The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Shiraz University of Medical Sciences [grant number: 27781].

Data availability The dataset used during the current study is not publicly available due to Shiraz University of Medical Sciences policy. Data dictionary and code pipeline are available at: Zenodo (DOI: 10.5281/zenodo.21915266), and the associated development repository is hosted on GitHub (<https://github.com/shashemi8888/deep-burns-survival-prediction.git>).

Declarations

Competing interests The authors declare no competing interests.

Ethics approval and consent to participate The study protocol was approved before the initiation of this study by the Ethics Committee of Shiraz University of Medical Sciences (approval ID: IR.SUMS.NUMIMG.REC.1402.032). The study was performed in compliance with this institutional guideline and the ethical guidelines for clinical research of the Iranian government. Participants gave their informed consent to participate.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.