Article

Machine Learning Models for Predicting In-Hospital Mortality in Burn Patients

Ann. Ital. Chir., 2025 96, 8: 1039–1046
https://doi.org/10.62713/aic.3944

Samet Şahin[1], Burak Yavuz[2], Onur Karaca[3], Merve Akın[4], Ali Emre Akgün[4]

[1]Department of General Surgery, Muğla Sıtkı Koçman University, 48000 Muğla, Turkey
[2]Department of General Surgery, Kozan State Hospital, 01510 Adana, Turkey
[3]Department of General Surgery, Koç University Hospital, 34010 İstanbul, Turkey
[4]Department of General Surgery, Bilkent City Hospital, 06800 Ankara, Turkey

AIM: To develop and evaluate predictive models for in-hospital mortality in burn patients using machine learning (ML) techniques.
METHODS: A retrospective cohort study was conducted using data from burn patients admitted to Ankara Bilkent City Hospital Burn Treatment Center between 2015 and 2020. Key variables including age, gender, total body surface area burned, burn depth, burn type, inhalation injury, inflammatory markers and inflammatory indexes were collected. Seven ML models—Logistic Regression, Random Forest, Support Vector Machine, Decision Tree, K-Nearest Neighbors, Naive Bayes, and Gradient Boosting—were trained and evaluated.
RESULTS: The cohort included 218 patients (mean age 42.5 ± 18.5 years; 69.7% male, 30.3% female), with an in-hospital mortality rate of 18.8% (n = 41). Logistic Regression had the best performance (accuracy: 88.6%, Receiver Operating Characteristic (ROC)-Area Under Curve (AUC): 0.906), while Random Forest achieved the highest accuracy (90.9%) and recall (97.2%). K-Nearest Neighbors excelled in recall (99.0%), Gradient Boosting balanced precision and recall (91.6% each, ROC-AUC: 0.744), and Support Vector Machine showed moderate results (accuracy: 84.0%, ROC-AUC: 0.864).
CONCLUSIONS: ML models, particularly Logistic Regression and Random Forest, demonstrated strong predictive capabilities for mortality in burn patients. This study supports the potential for ML in burn care, offering a data-driven approach for personalized prognosis and clinical decision-making. Further multicenter validation is recommended.

Keywords: burns/mortality; machine learning; risk assessment

Introduction

Burn injuries remain one of the most severe and life-threatening forms of trauma, with an estimated annual mortality of 180,000 cases worldwide according to the World Health Organization (WHO) [1]. Burn injuries remain one of the most complex and life-threatening types of trauma, necessitating timely and accurate assessment to inform clinical decision-making [2]. Mortality risk prediction is a crucial aspect of burn care, guiding treatment plans and resource allocation in critical settings. Traditionally, scores such as the Baux and revised Baux scores have been utilized to estimate mortality risk by considering factors like patient age, total body surface area (TBSA) burned, and the presence of inhalation injury [3–5]. Although these methods are practical and straightforward, they may be limited in scope, as they rely on a restricted set of variables and may not fully capture the nuanced clinical profile of each burn patient [6,7].

Recent advancements in machine learning (ML) have opened new avenues for enhancing predictive modeling in healthcare, offering a data-driven approach that can incorporate complex, multi-dimensional patient data. In contrast to traditional scores, which rely on fixed assumptions, ML algorithms discover patterns by analyzing the data directly. As demonstrated in other medical domains, such as oncology and sepsis prediction, ML has shown promise in improving outcome predictions and tailoring care approaches through more sophisticated analysis [8,9].

This study aims to evaluate the performance of various ML models—including Logistic Regression, Random Forest, Support Vector Machine, Decision Tree, K-Nearest Neighbors, Naive Bayes, and Gradient Boosting—in predicting in-hospital mortality among burn patients. By leveraging ML's ability to integrate complex, multi-dimensional data, we aim to develop tools capable of enhancing real-time clinical decision-making in modern intensive care unit (ICU) workflows.

Materials and Methods

This study was conducted to develop and evaluate predictive models for in-hospital mortality in burn patients using a retrospective cohort design. The study included all burn patients admitted and hospitalized in Ankara Bilkent City

Submitted: 2 January 2025 Revised: 14 March 2025 Accepted: 26 March 2025 Published: 14 July 2025
Correspondence to: Samet Şahin, Department of General Surgery, Muğla Sıtkı Koçman University, 48000 Muğla, Turkey (e-mail: samet-sahin@mu.edu.tr).

Hospital Burn Treatment Center between 2015 to 2020. Patients under 18 years old and patients with incomplete data were excluded. Data were extracted from the hospital's electronic health records, focusing on a comprehensive set of variables that could potentially influence mortality outcomes. This study was conducted in accordance with the principles outlined in the Declaration of Helsinki. Ethical approval was granted by the Ankara Bilkent City Hospital Institutional Ethical Review Board on 10 November 2021, with the approval number E2-21-927.

This retrospective study incorporated demographic variables (gender, age), burn injury characteristics (TBSA, burn depth, burn agent type, presence of inhalation injury), and admission laboratory values to predict in-hospital mortality. Laboratory parameters included C-reactive protein (CRP), delta neutrophil index (DNI), lymphocyte count, monocyte count, neutrophil count, lactate dehydrogenase (LDH), albumin, aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), and gamma-glutamyl transferase (GGT), all measured on admission. To further capture systemic inflammation and physiological stress, derived prognostic ratios were calculated, including neutrophil-to-lymphocyte ratio (NLR), lymphocyte-to-monocyte ratio (LMR), CRP-to-albumin ratio (CRP/Alb), DNI-to-albumin ratio (DNI/Alb), and CRP-to-lymphocyte ratio (CRP/Lymph). The binary outcome variable, in-hospital mortality, was defined as survival (0) or death (1). All variables were extracted from electronic health records, with no missing data due to institutional completeness protocols.

Seven machine learning algorithms were implemented: Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbors (KNN), Naive Bayes (NB), and Gradient Boosting (GB). These models were selected to encompass diverse algorithmic approaches, including linear, tree-based, and instance-based methods. Data preprocessing included standardization of features using StandardScaler to ensure comparability across scales. Mortality was stratified across training (60%), validation (20%), and test (20%) sets to preserve class distribution.

Hyperparameter optimization was conducted via grid search with 5-fold stratified cross-validation on the training set. Key tuned parameters included regularization strength (C) for LR, tree depth (max_depth) and ensemble size (n_estimators) for RF and GB, and kernel type for SVM. Model performance was evaluated on validation and test sets using accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (Receiver Operating Characteristic (ROC)-Area Under Curve (AUC)). Calibration curves (**Supplementary Fig. 1**) and decision curve analysis (**Supplementary Fig. 2**) were employed to assess probabilistic calibration and clinical utility, respectively.

Feature importance was analyzed through permutation importance for non-tree models (LR, SVM) and intrinsic Gini importance for tree-based algorithms (RF, GB, DT) (**Supplementary Data 1**). Statistical comparisons of cross-validated AUC-ROC scores were performed using DeLong test. All analyses were executed using Python (version 3.12, Python Software Foundation, Wilmington, DE, USA), leveraging pandas for data manipulation, scikit-learn for model implementation and evaluation, and matplotlib/seaborn for visualization. A p value lower than 0.05 was considered statistically significant.

Results

This study included 218 burn patients. Demographical and clinical information of the patients are presented in Table 1.

Table 1. Demographical and clinical data.

Parameter		Result
Age (years)		42.5 ± 18.5
Gender	Male	69.7% (n = 152)
	Female	30.3% (n = 66)
Total Length of Stay (Days)		15.5 (7–35)
Intensive Care Unit Length of Stay (Days)		7 (3–19)
Mortality		18.8% (n = 41)
Total Body Surface Area Burned (%)		16.50 (5.00–35.00)
Burn Depth	Superficial	1.4% (n = 3)
	Partial Thickness	45.4% (n = 99)
	Full Thickness	28.0% (n = 61)
	Deep	25.2% (n = 55)
Burn Agent	Flame	16.5% (n = 36)
	Arc	6.0% (n = 13)
	Scald Burn	33.5% (n = 73)
	Steam	8.7% (n = 19)
	Electrical	15.1% (n = 33)
	Chemical	9.6% (n = 21)
	Cold	1.4% (n = 3)
	Burning Agent	9.2% (n = 20)
Inhalation Burn		9.2% (n = 20)

Arc, Arc Flash. Categorical data is presented as percent (count). Numerical data is presented as mean ± standard deviation or median (interquartile range).

Laboratory data at admission are presented in Table 2.

The cross-validation results for the seven machine learning models are presented in Table 3. Validation set results are also provided in **Supplementary Data 2**. Among these, the Random Forest classifier achieved the highest overall classification accuracy (0.909), coupled with strong recall (0.972) and the top F1 Score (0.945). Logistic Regression, on the other hand, exhibited the most favorable discriminative performance as indicated by its highest ROC-AUC value (0.906), alongside robust accuracy (0.886), precision (0.942), recall (0.916), and F1 Score (0.929) (Table 3, Fig. 1).

Additionally, we evaluated model calibration and clinical utility across all classifiers. **Supplementary Fig. 1** presents the calibration curves based on the test set, com-

Table 2. Laboratory data.

Parameter	Value
CRP (mg/dL)	22.15 (2.75–82.75)
Lymphocyte ($\times 10^9$/L)	1.45 (1.06–2.12)
Monocyte ($\times 10^9$/L)	0.64 (0.45–0.88)
Neutrophil ($\times 10^9$/L)	9.14 (5.93–14.88)
LDH (U/L)	300.50 (227.75–456.25)
Albumin (g/L)	37.50 (31.00–43.00)
AST (U/L)	32.00 (21.75–53.00)
ALT (U/L)	27.00 (19.00–40.25)
ALP (U/L)	72.00 (57.75–94.00)
GGT (U/L)	22.00 (14.00–38.25)
CRP to Albumin	0.62 (0.06–2.35)
DNI to Albumin	0.0033 (0.0024–0.0656)
CRP to Lymphocyte	14.29 (0.92–56.67)
LMR	2.48 (1.45–3.70)
NLR	6.10 (3.45–11.33)
DNI (%)	0.10 (0.10–2.03)

Numerical data is presented as median (interquartile range). CRP, C-reactive protein; LDH, lactate dehydrogenase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALP, alkaline phosphatase; GGT, gamma-glutamyl transferase; DNI, delta neutrophil index; LMR, lymphocyte-to-monocyte ratio; NLR, neutrophil-to-lymphocyte ratio.

paring the mean predicted probabilities against the fraction of positive outcomes. Most models showed reasonable calibration, with Logistic Regression and Support Vector Machine demonstrating the closest adherence to ideal calibration across the prediction range. **Supplementary Fig. 2** displays the Decision Curve Analysis (DCA) based on the test set, illustrating the net benefit of each model across different threshold probabilities. The majority of models maintained positive net benefit values within the lower threshold ranges (0.0–0.4).

Support Vector Machine demonstrated balanced metrics (accuracy = 0.840, precision = 0.914, recall = 0.888, F1 Score = 0.901, ROC-AUC = 0.864), whereas the Decision Tree classifier showed notable recall (0.944) and precision (0.918), resulting in a strong F1 Score (0.931). However, its ROC-AUC (0.784) was comparatively lower. K-Nearest Neighbors achieved the highest recall (0.990), while also maintaining moderate accuracy (0.863) and a solid F1 Score (0.923). Naive Bayes performed adequately, with an accuracy of 0.840, precision of 0.939, recall of 0.861, F1 Score of 0.898, and ROC-AUC of 0.857. Lastly, Gradient Boosting demonstrated balanced accuracy, precision, and recall (0.863, 0.916, and 0.916, respectively), though it exhibited a relatively modest ROC-AUC of 0.744. Confusion matrices are presented in Fig. 2.

Discussion

Burn injuries represent a complex clinical challenge, with mortality risk influenced by dynamic interactions between systemic inflammation, physiological reserve, and injury severity. This retrospective cohort study of 218 adult burn patients (2015–2020) leverages ML to integrate multidimensional predictors—including demographic factors, burn characteristics, admission laboratory profiles, and novel inflammatory ratios—into prognostic models for in-hospital mortality. By evaluating seven ML algorithms, we demonstrate that Logistic Regression (ROC-AUC: 0.906) and Random Forest (accuracy: 0.909, recall: 0.972) achieved robust predictive performance, outperforming conventional metrics in sensitivity while maintaining clinical interpretability.

Our results align with prior study such as Yeh *et al*. [10], who investigated AI and ML models in burn patients to predict adverse outcomes, including the need for graft surgery and prolonged hospital stays. Yeh *et al*. [10] found that Random Forest provided the highest AUC (81.1% for prolonged hospital stay and 78.8% for graft requirement), while extreme Gradient Boosting (XGBoost) performed best for adverse outcomes (AUC 87.2%). The divergence in model performance between their work and ours—particularly the lower AUCs in their study—may reflect fundamental differences in outcome heterogeneity. Their simultaneous evaluation of multiple endpoints (grafting, prolonged stay) likely introduced competing risks and diluted model specificity, whereas our focus on mortality as a singular endpoint allowed for optimized feature alignment. Similarly, Park *et al*.'s [11] work on critically ill burn patients demonstrated strong predictive accuracy for Random Forest, achieving an AUC of 0.922 for 90-day mortality, reinforcing the value of Random Forest in our study as well, where it produced the highest ROC-AUC among models, indicating robust predictive power. While they reported a higher AUC for Random Forest (0.922 vs. 0.795 in our study), this divergence likely reflects critical differences in cohort severity and outcome definitions. Their cohort exclusively analyzed critically ill surgical patients with substantially larger burns (mean TBSA 38.5% in survivors vs. 16.5% in our cohort) and a higher prevalence of inhalation injuries (31% vs. 9.2% in our study), factors that inherently amplify the predictive weight of variables like TBSA and American Society of Anesthesiologists Physical Status (ASA-PS) in their models. Furthermore, their focus on 90-day mortality—capturing delayed postoperative complications—contrasts with our in-hospital mortality endpoint, which prioritizes acute physiological derangements reflected in admission inflammatory markers. This distinction is evident in feature importance: their RF model emphasized perioperative factors (TBSA, Red Blood Cell Distribution Width (RDW), ASA-PS), whereas ours highlighted inflammatory ratios (NLR, CRP/albumin), suggesting that outcome timing shapes predictive variable relevance. The superior AUC of their RF model may also stem from their larger sample size (n = 731 vs. 218), which better accommodates tree-based algorithms' need for data depth to model complex interactions in high-severity cohorts.

Table 3. Model metrics of the test set.

Model	Accuracy	Precision	Recall	F1 Score	ROC-AUC
Logistic Regression	0.886	0.942	0.916	0.929	0.906^{a}
Random Forest	0.909	0.921	0.972	0.945	0.795
Support Vector Machine	0.840	0.914	0.888	0.901	0.864
Decision Tree	0.886	0.918	0.944	0.931	0.784
K-Nearest Neighbors	0.863	0.857	0.990	0.923	0.817
Naive Bayes	0.840	0.939	0.861	0.898	0.857
Gradient Boosting	0.863	0.916	0.916	0.916	0.744

^{a}: $p < 0.05$ compared to Decision Tree. ROC, Receiver Operating Characteristic; AUC, Area Under Curve.

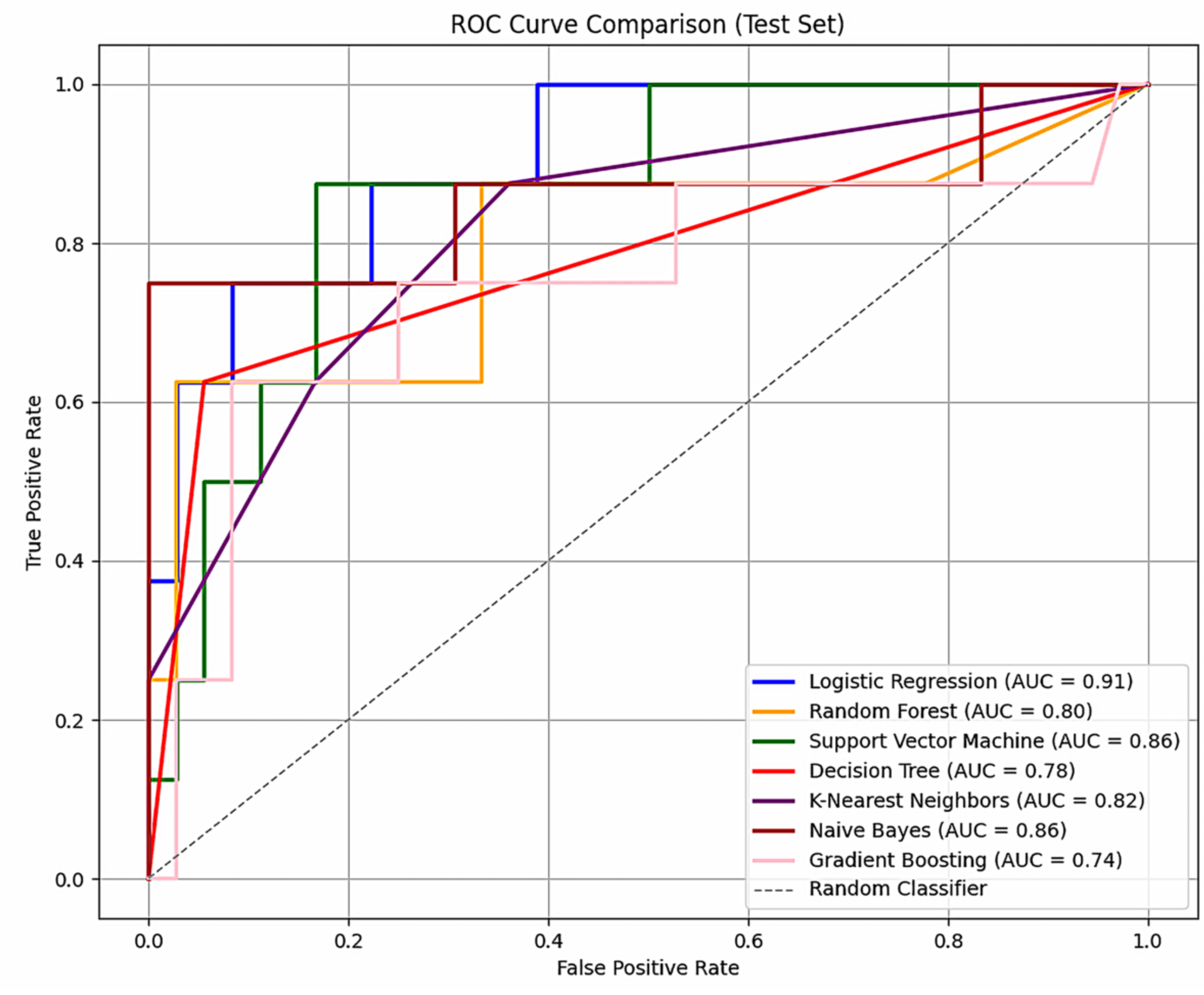

Fig. 1. Receiver Operating Characteristic curves of models. ROC, Receiver Operating Characteristic; AUC, Area Under Curve.

In a study by Yazıcı *et al*. [12], six ML algorithms were assessed for predicting burn-related mortality based on risk factors, including age, gender, TBSA, full-thickness burns, and inhalation injury. Their findings underscored AdaBoost as the top performer, achieving 90% accuracy and an AUC of 92%, emphasizing age and TBSA as key predictors [12]. Çinar *et al*. [13] also conducted a study involving 1064 patients hospitalized in a burn center between 2016 and 2022 to predict mortality risk using machine learning models. They analyzed 40 parameters, including demographic and biochemical data, and employed various machine learning methods, with artificial neural networks (ANNs) showing the highest accuracy (95.92%) in predicting outcomes. Their study highlighted the potential of machine learning in clinical decision support systems, emphasizing its ability to enhance early and accurate mortality risk assessment in burn patients [13]. Our findings showed high recall in SVM and K-Nearest Neighbors,

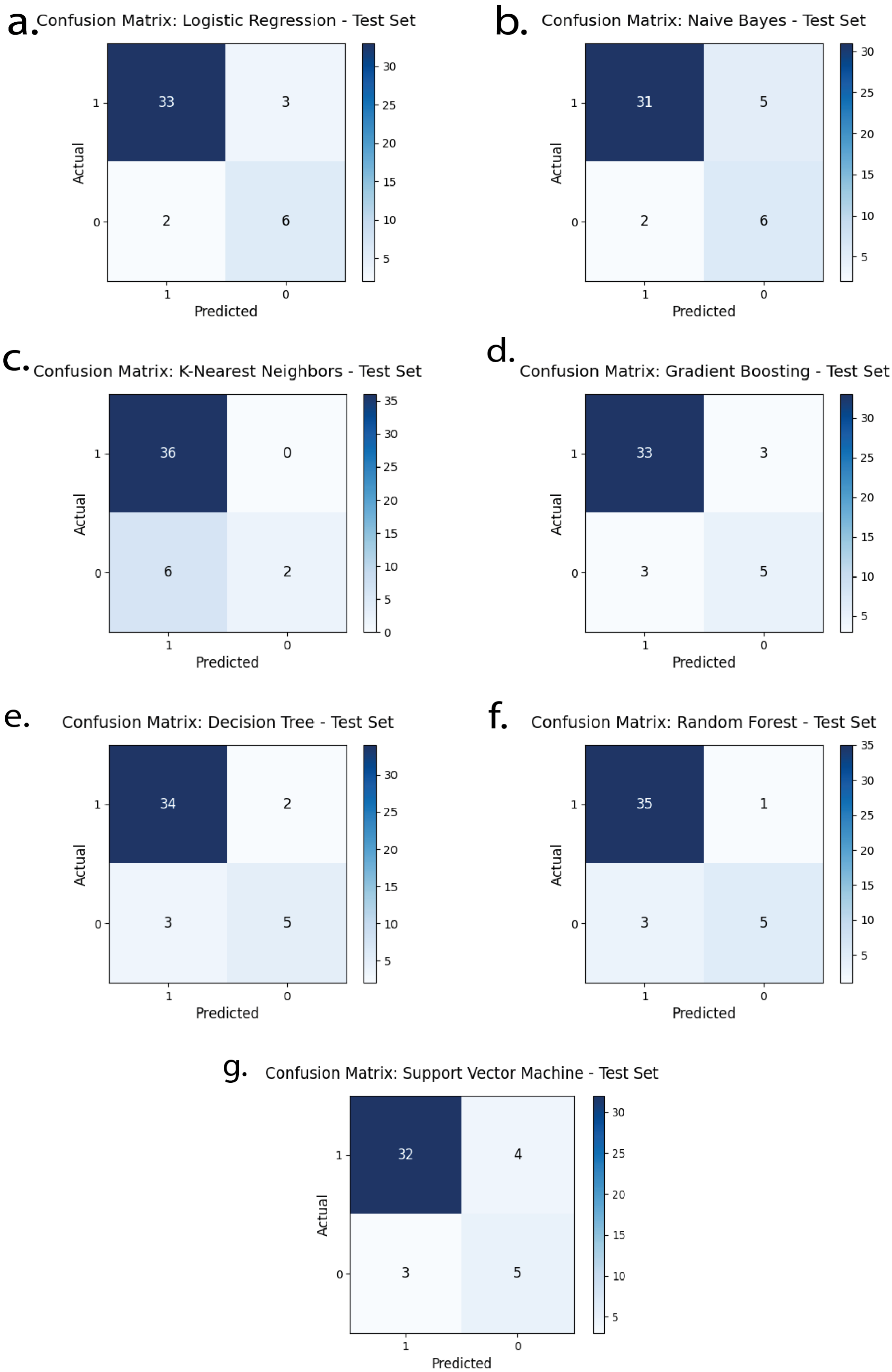

Fig. 2. Confusion matrices for various machine learning models predicting in-hospital mortality in burn patients. (a) Logistic Regression, (b) Naive Bayes, (c) K-Nearest Neighbor, (d) Gradient Boosting, (e) Decision Tree, (f) Random Forest, (g) Support Vector Machine (SVM).

though with comparatively lower precision and F1 scores. This difference may stem from our use of indices tied to inflammatory markers, a focus that diverges from traditional predictive parameters and possibly limits precision.

Comparing our findings with Stylianou *et al.* [14], who used various ML models such as ANN, RF, and LR for burn mortality prediction, we observed similarities and differences. Stylianou *et al.*'s study [14] reported ANN as the best performer for AUC, while RF had a high positive predictive value despite lower sensitivity. Their ANN outperformed our Logistic Regression in AUC (0.971 vs. 0.906), but this stark contrast arises from fundamental differences in dataset scale and clinical context. Their analysis of 66,611 patients with a mortality rate of 1.27%—representing a general burn population—contrasts sharply with our cohort's higher acuity (18.8% mortality, median TBSA 16.5% vs. their 1.5%). The extreme class imbalance in their data (non-fatal scalds comprising 41% of cases) inherently favors ANN's capacity to detect subtle nonlinear patterns in low-risk populations, whereas our smaller, high-severity cohort (predominantly flame/electrical burns) prioritized acute inflammatory derangements captured effectively by interpretable linear models. Furthermore, their focus on admission-to-discharge mortality prediction using static variables (age, TBSA, inhalation injury) differs from our inclusion of dynamic inflammatory ratios, which may explain RF's superior sensitivity in our setting. Their optimization for maximum Youden's index versus our emphasis on recall reflects divergent clinical priorities: screening low-risk populations versus avoiding missed predictions in critically ill patients.

Sepsis is also a key factor influencing mortality in burn patients, with conventional sepsis indicators often underperforming in burn cases. Tran *et al.* [15] developed a "machine intelligence learning optimizer" (MILO) for predicting sepsis in burn patients, comparing it with non-automated ML methods. In their retrospective analysis of 211 patients, the MILO approach using K-Nearest Neighbors yielded an impressive 90% accuracy and a ROC-AUC of 0.96, underscoring the potential of advanced ML models for predicting complications like sepsis with high accuracy [15].

Another relevant study by Rashidi *et al.* [16] focused on early acute kidney injury (AKI) prediction in burn and trauma patients using biomarkers (e.g., neutrophil gelatinase-associated lipocalin (NGAL), NT-proBNP, urine output and creatinine). Their ML models showed that NGAL, combined with NT-proBNP or creatinine, allowed AKI prediction up to 61.8 hours earlier than standard criteria. With ML performance improvements, predictive models using accessible, routinely monitored markers may become more effective in clinical practice [16].

In a recent systematic review by Taib *et al.* [17], machine learning models demonstrated sensitivities and specificities of 92.9% and 93.4%, respectively, for mortality prediction in burn patients, surpassing the modified Baux score and using readily available data. This review reinforces the clinical applicability and predictive strength of ML approaches in managing burn-related mortality [17].

In the literature, ML models demonstrated superior predictive performance compared to the Baux score. In a study conducted by Maxwell *et al.* [18] on 100 patients, the ROC-AUC of the Baux score in predicting mortality was found to be 0.682 ($p < 0.05$). Similarly, a 2020 study by Choi *et al.* [19] involving 183 burn patients reported a revised Baux score ROC-AUC of 0.840 for predicting mortality (95% CI: 0.76–0.91, $p < 0.001$). Fransén *et al.* [20] explored the application of machine learning (ML) algorithms for predicting mortality in burn patients, benchmarking their performance against the established Baux and revised Baux scores. Using data from 92 patients, ML models such as Extreme Boosting, Random Forest, and Support Vector Machine (SVM) achieved an AUC of 0.920, comparable to the Baux scores, which reached 0.850 and 0.840. No statistically significant difference was found between the ML models and the Baux scores, suggesting that ML could offer similar predictive accuracy within clinical settings. We intentionally adopted an ML-based methodology to explore the potential for advanced algorithms to match or surpass conventional prediction methods, while also broadening the scope of predictive variables beyond age and TBSA. This approach allowed us to assess the adaptability and robustness of ML models in real-world clinical data, potentially uncovering more nuanced insights into mortality risk factors. Unlike the Baux score, which is limited to age, TBSA, and inhalation injury, our models and the models in the literature incorporated additional clinical and laboratory parameters, potentially providing a more nuanced and individualized risk assessment.

Although the findings support the feasibility of ML for mortality prediction, limitations—such as the small sample size and single-center data—highlight the need for further research with larger, multicenter studies to validate these results across diverse clinical contexts. This work reflects an ongoing effort to advance predictive modeling in burn care, where complex models may pave the way for personalized prognostic tools that can be tailored to modern ICU data and patient variables. While the models performed well overall, their applicability to specific subgroups, such as patients with extensive TBSA burns or severe inhalation injuries, requires further investigation. These subgroups often present unique challenges, including higher mortality risks and atypical clinical trajectories, which may influence model performance. Stratified analyses in future studies could elucidate subgroup-specific model efficacy. Future research should validate these models using multi-center datasets to ensure broader applicability across diverse clinical settings.

Conclusions

Our study supports the view that machine learning models, particularly Logistic Regression and Random Forest, provide robust predictive capabilities for mortality in burn patients. These models have the potential to enhance clinical decision-making by providing real-time risk assessments that could guide resource allocation and treatment prioritization in intensive care units.

To ensure broader applicability, future efforts should focus on developing tools tailored for specific patient subgroups, such as those with extensive burns, inhalation injuries, or significant comorbidities. Integrating these models into electronic health record systems could facilitate real-time mortality predictions, enabling personalized treatment strategies and potentially improving outcomes for high-risk patients. Further multi-center validation studies are essential to confirm the generalizability of these findings and their clinical impact.

Availability of Data and Materials

The data analyzed are available from the corresponding author upon reasonable request.

Author Contributions

SŞ, MA and AEA designed the research study; SŞ, BY, OK, and AEA performed the research; BY and OK collected and analyzed the data. SŞ and MA were involved in drafting the manuscript, and all authors were involved in revising it critically for important intellectual content. All authors have given final approval of the version to be published. All authors have participated sufficiently in the work to take public responsibility for appropriate portions of the content and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Written informed consent was obtained from all participants. This study was conducted in accordance with the principles outlined in the Declaration of Helsinki. Ethical approval was granted by the Ankara Bilkent City Hospital Institutional Ethical Review Board on 10 November 2021, with the approval number E2-21-927.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.62713/aic.3944.

References

[1] Smolle C, Cambiaso-Daniel J, Forbes AA, Wurzer P, Hundeshagen G, Branski LK, *et al*. Recent trends in burn epidemiology worldwide: A systematic review. Burns: Journal of the International Society for Burn Injuries. 2017; 43: 249–257. https://doi.org/10.1016/j.burns.2016.08.013.

[2] Dobson GP, Morris JL, Letson HL. Pathophysiology of Severe Burn Injuries: New Therapeutic Opportunities From a Systems Perspective. Journal of Burn Care & Research: Official Publication of the American Burn Association. 2024; 45: 1041–1050. https://doi.org/10.1093/jbcr/irae049.

[3] Osler T, Glance LG, Hosmer DW. Simplified estimates of the probability of death after burn injuries: extending and updating the baux score. The Journal of Trauma. 2010; 68: 690–697. https://doi.org/10.1097/TA.0b013e3181c453b3.

[4] Edgar MC, Bond SM, Jiang SH, Scharf IM, Bejarano G, Vrouwe SQ. The Revised Baux Score as a Predictor of Burn Mortality: A Systematic Review and Meta-Analysis. Journal of Burn Care & Research: Official Publication of the American Burn Association. 2023; 44: 1278–1288. https://doi.org/10.1093/jbcr/irad075.

[5] Roberts G, Lloyd M, Parker M, Martin R, Philp B, Shelley O, *et al*. The Baux score is dead. Long live the Baux score: a 27-year retrospective cohort study of mortality at a regional burns service. The Journal of Trauma and Acute Care Surgery. 2012; 72: 251–256. https://doi.org/10.1097/TA.0b013e31824052bb.

[6] Meyfroidt G, Güiza F, Ramon J, Bruynooghe M. Machine learning techniques to examine large patient databases. Best Practice & Research. Clinical Anaesthesiology. 2009; 23: 127–143. https://doi.org/10.1016/j.bpa.2008.09.003.

[7] Bica I, Alaa AM, Lambert C, van der Schaar M. From Real-World Patient Data to Individualized Treatment Effects Using Machine Learning: Current and Future Methods to Address Underlying Challenges. Clinical Pharmacology and Therapeutics. 2021; 109: 87–100. https://doi.org/10.1002/cpt.1907.

[8] Islam MM, Nasrin T, Walther BA, Wu CC, Yang HC, Li YC. Prediction of sepsis patients using machine learning approach: A meta-analysis. Computer Methods and Programs in Biomedicine. 2019; 170: 1–9. https://doi.org/10.1016/j.cmpb.2018.12.027.

[9] Yala A, Mikhael PG, Strand F, Lin G, Smith K, Wan YL, *et al*. Toward robust mammography-based models for breast cancer risk. Science Translational Medicine. 2021; 13: eaba4373. https://doi.org/10.1126/scitranslmed.aba4373.

[10] Yeh CC, Lin YS, Chen CC, Liu CF. Implementing AI Models for Prognostic Predictions in High-Risk Burn Patients. Diagnostics (Basel, Switzerland). 2023; 13: 2984. https://doi.org/10.3390/diagnostics13182984.

[11] Park JH, Cho Y, Shin D, Choi SS. Prediction of Mortality after Burn Surgery in Critically Ill Burn Patients Using Machine Learning Models. Journal of Personalized Medicine. 2022; 12: 1293. https://doi.org/10.3390/jpm12081293.

[12] Yazıcı H, Uçar AD, Namdaroğlu O, Yıldırım M. Mortality prediction models for severe burn patients: Which one is the best? Ulusal Travma Ve Acil Cerrahi Dergisi = Turkish Journal of Trauma & Emergency Surgery: TJTES. 2022; 28: 790–795. https://doi.org/10.14744/tjtes.2021.29540.

[13] Çinar MA, Ölmez E, Erkiliç A, Bayramlar K, Er O. Machine learning models for early prediction of mortality risk in patients with burns: A single center experience. Journal of Plastic, Reconstructive & Aesthetic Surgery: JPRAS. 2024; 89: 14–20. https://doi.org/10.1016/j.bjps.2023.11.048.

[14] Stylianou N, Akbarov A, Kontopantelis E, Buchan I, Dunn KW. Mortality risk prediction in burn injury: Comparison of logistic regression with machine learning approaches. Burns: Journal of the International Society for Burn Injuries. 2015; 41: 925–934. https://doi.org/10.1016/j.burns.2015.03.016.

[15] Tran NK, Albahra S, Pham TN, Holmes JH, 4th, Greenhalgh D, Palmieri TL, *et al*. Novel application of an automated-machine learn-

ing development tool for predicting burn sepsis: proof of concept. Scientific Reports. 2020; 10: 12354. https://doi.org/10.1038/s41598-020-69433-w.

[16] Rashidi HH, Sen S, Palmieri TL, Blackmon T, Wajda J, Tran NK. Early Recognition of Burn- and Trauma-Related Acute Kidney Injury: A Pilot Comparison of Machine Learning Techniques. Scientific Reports. 2020; 10: 205. https://doi.org/10.1038/s41598-019-57083-6.

[17] Taib BG, Karwath A, Wensley K, Minku L, Gkoutos GV, Moiemen N. Artificial intelligence in the management and treatment of burns: A systematic review and meta-analyses. Journal of Plastic, Reconstructive & Aesthetic Surgery: JPRAS. 2023; 77: 133–161. https://doi.org/10.1016/j.bjps.2022.11.049.

[18] Maxwell D, Rhee P, Drake M, Hodge J, Ingram W, Williams R. Development of the Burn Frailty Index: A prognostication index for elderly patients sustaining burn injuries. American Journal of Surgery. 2019; 218: 87–94. https://doi.org/10.1016/j.amjsurg.2018.11.012.

[19] Choi KJ, Pham CH, Collier ZJ, Mert M, Ota RK, Li R, *et al*. The Predictive Capacity of American Society of Anesthesiologists Physical Status (ASA PS) Score in Burn Patients. Journal of Burn Care & Research: Official Publication of the American Burn Association. 2020; 41: 803–808. https://doi.org/10.1093/jbcr/iraa060.

[20] Fransén J, Lundin J, Fredén F, Huss F. A proof-of-concept study on mortality prediction with machine learning algorithms using burn intensive care data. Scars, Burns & Healing. 2022; 8: 20595131211066585. https://doi.org/10.1177/20595131211066585.