
Admission-Only ML Supports Sepsis Triage in Burn ICU



Sepsis remains the leading cause of death in burn care, yet early recognition is complicated by persistent hyperinflammatory responses and altered baseline physiology following major injury. A streamlined machine learning approach has been developed to identify patients at risk of sepsis using information available at admission to the intensive care unit (ICU). Trained on 6629 adult cases from 11 centres contributing to the German burn registry, the model relies on six routinely captured variables and reaches high discriminative performance. By offering immediate risk stratification without post-admission data, the approach is positioned for practical integration into ICU workflows where timely decision-making and resource allocation are critical.

Large Registry Enables Lean Predictive Approach

The cohort comprised 6629 patients, of whom 521 (7.9%) developed sepsis during hospitalisation. The analysis focused explicitly on admission-level information to avoid dependence on dynamic trends that emerge hours to days before onset. The six variables used were age, burned body surface area (TBSA), deep partial-thickness burns (burn depth 2b), full-thickness burns (burn depth 3), inhalation injury and hypertension. These features were consistently identified through multiple selection techniques, including regularisation and recursive elimination and reflect established components of initial burn assessment.

Must Read: [Avoiding Bias in ICU Prognostic Models](#)

Sepsis cases differed from non-sepsis cases in several baseline characteristics captured in the registry. Those who developed sepsis were older, had larger TBSA and deeper burns and more often presented with inhalation injury. Comorbidities such as diabetes mellitus, coronary heart disease and hypertension were more common in the sepsis group. Clinical course markers associated with severity, such as ventilation and pneumonia, were also more frequent among patients who became septic and mortality was higher in this group. While such variables underscore risk, the modelling strategy deliberately excluded post-admission outcomes and potential leakage features to maintain a clean admission-only predictor.

Feature set construction compared four configurations: an eight-feature Intersection set selected by all automated methods, a 12-feature High Frequency set chosen by most methods, a six-feature hypothesis-driven set based on core clinical variables and a Minimalistic set removing hypertension and inhalation injury to test feasibility with only four inputs. This staged design allowed evaluation of the trade-off between simplicity and performance, with particular emphasis on clinical deployability at the point of admission.

Performance Balances Sensitivity and Practicality

Across Random Forest, Logistic Regression, LightGBM and XGBoost algorithms, the six-feature configuration delivered the strongest overall performance. A Random Forest trained on these variables achieved an area under the receiver operating characteristic curve (AUROC) of 0.91 with sensitivity 0.81, specificity 0.85 and negative predictive value (NPV) 0.987. Logistic Regression performed similarly with AUROC 0.90, sensitivity 0.81 and specificity 0.85. Expanding to 12 features yielded only marginal changes, indicating diminishing returns beyond the core subset. Even the four-feature Minimalistic variant maintained competitive performance with AUROC 0.90, highlighting robustness where rapid or resource-limited integration is needed.

The modelling pipeline prioritised recall to minimise false negatives given the clinical risk of missed cases. This emphasis led to a modest positive predictive value (PPV) around 0.31–0.35 in several configurations, a typical challenge for low-prevalence outcomes. Nonetheless, the consistently high NPV, reaching 0.98, indicates strong reliability for identifying patients at low risk, supporting cautious restraint in initiating therapies that carry harm if used indiscriminately.

Receiver operating characteristic and precision–recall analyses with confidence intervals confirmed stable discrimination across algorithms and feature sets. The final model was selected on the basis of balanced accuracy, interpretability and admission-time feasibility. Internal validation used five-fold stratified cross-validation and a held-out test split with preprocessing steps to handle missingness according to data type and clinical reasoning. The outcome definition aligned with established burn-specific consensus criteria.

Interpretability, Clinical Use and Limitations

Model explainability assessed using SHAP identified TBSA as the most influential contributor followed by full-thickness burns and age. Partial dependence analyses indicated a steep rise in predicted risk at TBSA around 20–30% with a ceiling effect near 40–50%, a strong contribution of full-thickness injury even at lower increments and a non-linear age effect with greater impact beyond approximately 40 years. Inhalation injury and hypertension showed step-like influences consistent with their binary nature. These patterns align with recognised clinical factors while providing quantitative insights into relative contributions at admission.

The intended use is early risk stratification, not real-time diagnosis at the moment sepsis is suspected. A high-risk flag should not prompt immediate antimicrobial therapy given the modest PPV. Rather, it can justify heightened surveillance, lower thresholds for targeted diagnostics and consideration of advanced monitoring where local practice varies. Conversely, the high NPV offers a firmer basis to defer antibiotics when inflammatory signs are ambiguous, supporting stewardship while maintaining vigilance. Post-hoc inspection of predictions suggested that flagged patients who did not develop sepsis still represented a clinically relevant intermediate-risk profile although therapeutic decisions should remain grounded in comprehensive assessment.

Several limitations temper generalisation. Performance has been evaluated only within the registry cohort and external validation in other settings or regions is pending. Variation in sepsis definitions for burn patients persists, the modelling here used American Burn Association criteria and metrics may differ where alternative frameworks are applied. Despite these constraints, the reliance on routine admission variables makes the approach straightforward to test prospectively. Plans are in progress to make the model and documentation accessible to enable broader validation and implementation.

A lean, admission-only model using six routine variables identifies burn patients at risk of sepsis with high discrimination and very strong ability to rule out low-risk cases. By avoiding dependence on post-admission trends, the tool suits immediate ICU risk stratification, informing surveillance intensity and stewardship decisions without dictating treatment. Pending external validation and local calibration, its combination of accuracy, interpretability and operational simplicity presents a practical route to strengthen early sepsis risk management in burn care.

Source: [npj Digital Medicine](#)

Image Credit: [iStock](#)

Published on : Tue, 28 Oct 2025