

# Identification of Prognostic Clinical Features in Grade 4 Immune-Related Adverse Events: A Triangulation Study

Yulin Wu<sup>1-3,\*</sup>, Junyao Chen<sup>1,\*</sup>, Sijia Tian<sup>1</sup>, Zhaojie Lin<sup>1</sup>, Yong Li<sup>1</sup>, Cuihan Wang<sup>4</sup>, Qianying Lu<sup>1</sup>, Lu Lu<sup>1</sup>, Yanmei Zhao<sup>1</sup>

<sup>1</sup>School of Disaster and Emergency Medicine, Tianjin University, Tianjin, 300072, People's Republic of China; <sup>2</sup>Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin, 300060, People's Republic of China; <sup>3</sup>Tianjin's Clinical Research Center for Cancer, Tianjin, 300060, People's Republic of China; <sup>4</sup>Department of Infectious Disease, Tianjin Hospital of Integration of Traditional Chinese and Western Medicine & Nankai Hospital, Tianjin, 300102, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Yanmei Zhao; Lu Lu, School of Disaster and Emergency Medicine, Tianjin University, No. 92 Weijin Road, Tianjin, 300072, People's Republic of China, Tel/Fax +86-22-27893596, Email zhaoyanmei@126.com; lulu\_998543@tju.edu.cn

**Background:** Grade 4 immune-related adverse events (irAEs) is life-threatening complications of immune checkpoint inhibitor therapy. Due to its rarity and scarcity of data, there is a lack of systematic research on key factors influencing poor prognosis. This exploratory study aimed to identify clinical features robustly associated with mortality in patients with grade 4 irAEs. Given the extremely small sample size and high-dimensional data, a triangulation approach integrating traditional univariate statistics and machine learning to maximize the reliability of feature selection.

**Methods:** This study included 26 cancer patients admitted to the ICU for grade 4 irAEs. To maximize robustness from limited data, a “triangulation” approach was employed. Prognostic features were independently identified through two parallel approaches: (1) traditional univariate statistical analysis, and (2) multiple machine learning algorithms evaluated by Leave-One-out Cross-Validation. Features consistently highlighted as significant by both independent methodologies were integrated to form a final high-confidence feature set.

**Results:** Univariate analysis identified 21 features significantly associated with mortality. Machine learning analysis refined this to 11 important features. Through “triangulation”, 8 features were consistently validated: body mass index and VEGF-inhibitors were inversely associated with mortality, while vasopressor therapy, oxygen therapy, lactate levels at day 1 and 2, pneumonia and percentage of neutrophils, exhibited a positive correlation with the mortality.

**Conclusion:** This small-sample exploratory study identified 8 routinely available early ICU clinical features robustly associated with mortality in grade 4 irAEs patients using a “triangulation” framework. These characteristics highlight the pivotal roles of shock, respiratory failure, and inflammation. While not directly constructing a clinical prediction model, they may facilitate early risk stratification and provide hypotheses for prioritized validation in future large-sample studies.

**Keywords:** grade 4 immune-related adverse events, irAEs, intensive care unit, ICU, machine learning, triangulation, prognostic factors

## Introduction

Over the past decade, the field of cancer therapy has witnessed a profound transformation, driven primarily by the advent and widespread application of immune checkpoint inhibitors (ICIs).<sup>1</sup> These agents work by reactivating the patient's own immune system to recognize and attack cancer cells, fundamentally altering the treatment paradigm for various advanced or metastatic cancers—an achievement recognized by the Nobel Prize in Medicine in 2018.<sup>2</sup> ICIs have demonstrated promising efficacy against several malignancies previously considered difficult to treat, such as advanced non-small cell lung cancer,<sup>3</sup> metastatic renal cell carcinoma,<sup>4</sup> and malignant melanoma.<sup>5</sup> Consequently, ICIs have gradually become a mainstay in cancer treatment in recent years,<sup>6</sup> with PD-1/PD-L1 inhibitors being the most widely used.<sup>7</sup> However, the use of ICIs has introduced a novel

spectrum of side effects known as immune-related adverse events (irAEs), posing significant challenges for both patients and clinicians.<sup>8</sup> Unlike traditional chemotherapy or molecularly targeted therapies, irAEs are difficult to predict, can affect nearly any organ system, and may occur at any time during treatment or even after its completion.<sup>9</sup> According to the Common Terminology Criteria for Adverse Events (CTCAE) established by the US National Cancer Institute, irAEs are graded from 1 (mild/asymptomatic, requiring no intervention) to 5 (death).<sup>10</sup> In recent years, multiple large-scale clinical trials and real-world studies have characterized the general epidemiology of irAEs. According to the available literature, the incidence of irAEs among patients receiving ICIs ranges from 60% to 90%,<sup>11</sup> while the incidence of grade  $\geq 3$  severe irAEs ranges from approximately 7%–20%.<sup>12,13</sup> Several studies suggest that patients experiencing irAEs may have better oncological outcomes.<sup>14,15</sup> However, the occurrence of severe irAEs (Grade  $\geq 3$ ) not only can lead directly to patient death<sup>16</sup> but also demands greater clinical vigilance and intervention. Approximately 0.3–1.3% of patients die from irAEs.<sup>17</sup> These severe adverse events not only directly impact short-term patient survival but also lead to ICI treatment interruption, complications from prolonged immunosuppressive therapy, and significantly reduced quality of life. Therefore, identifying and predicting the clinical outcomes, particularly the final survival status, of patients with severe irAEs has become a critical unresolved issue in clinical practice.<sup>18</sup>

Currently, some studies have attempted to explore prognostic factors for irAEs based on clinical indicators. However, most combine irAEs of different grades for analysis,<sup>19</sup> an approach that may obscure the differential impacts associated with varying severity levels. Recently, some studies have begun to explore predictors of severe irAEs (grade  $\geq 3$ ). For instance, Acar conducted a retrospective cohort study of 593 patients with solid tumors and found that baseline eosinophil count, baseline red cell distribution width (RDW) were significantly associated with an increased risk of severe irAEs.<sup>20</sup> These inexpensive and readily available routine blood markers offer new insights for early identification of high-risk patients. However, the endpoint of that study was grade  $\geq 3$  irAEs without a separate analysis of the most critical grade 4 irAEs, and most patients were not admitted to the ICU. Also, the study focused more on risk factors for the development of irAEs rather than on the final prognostic outcomes.

As can be seen, research focusing on the survival outcomes of patients with Grade 4 (life-threatening) irAEs remains scarce. This gap primarily stems from the low incidence of Grade 4 irAEs, limited case numbers, and difficulties in data acquisition,<sup>21</sup> leaving this patient population in a “data desert” and hindering prospective clinical trials or large-scale cohort studies. Due to the sample size limitation, this study did not aim to construct a clinical prediction model with generalization ability, but rather focused on a more fundamental prior question: which clinical characteristics may be associated with mortality outcomes among very critical patients admitted to ICUs for grade 4 irAEs.

The pathogenesis of irAEs involves complex multi-system, multi-factor interactions, often characterized by non-linear relationships among variables. Traditional statistical methods often have limited power when handling such high-dimensional, small-sample data,<sup>22</sup> struggling to identify key predictive factors and build robust prediction models. Meanwhile, machine learning methods, leveraging their strengths in modeling complex feature interactions, non-linear relationships, and high-dimensional data, are gradually being introduced into medical predictive modeling.<sup>23</sup>

However, when the sample size is very small and there are many candidate variables, both traditional statistical methods and machine learning approaches have inherent limitations when used alone. Traditional univariate analysis, while controlling for type I error, cannot capture complex interactions or nonlinear relationships among variables and may miss true associations in small samples. Machine learning methods, although adept at identifying complex patterns in high-dimensional data, are highly prone to overfitting when the sample size is much smaller than the number of features, leading to spurious generalizability. To address this challenge, we introduced a “triangulation” approach as our core analytical strategy. Triangulation originates from social sciences and epidemiology, based on the principle that when multiple independent analytical pathways with different methodological assumptions converge on the same finding, the likelihood of a true association increases substantially, while method-specific biases may cancel each other out. The goal of this design is not to pursue predictive accuracy in the traditional sense, but rather to maximize the robustness and reproducibility of feature selection under limited data conditions, thereby providing focused hypotheses for subsequent large-scale validation.

Based on the above background, this study aims to conduct an exploratory preliminary investigation to gain an initial understanding of the characteristics of this high-risk population, thereby enriching our knowledge of this group. The primary objective is not to construct a clinical prediction model, but rather to identify clinical features within the data that may be robustly associated with mortality, providing hypotheses for prioritized validation in future large-scale studies. In

this study, we collected demographic information, treatment history, past history and biochemical test indexes of patients who needed to be admitted to the ICU due to grade 4 irAEs in the Intensive Care Unit of Tianjin Medical University Cancer Hospital between 2018 and 2024. The key predictors of final prognostic outcomes in this understudied population were explored by traditional statistical methods and machine learning methods, respectively. The main contributions of this study include:

1. This study focuses on the survival outcomes of patients with Grade 4 immune-related adverse reactions to conduct a predictive exploration, providing preliminary clues and hypotheses in this field;
2. Propose and apply the “triangulation” analysis framework, which combines traditional statistics and machine learning, to realize exploratory studies on small clinical samples;
3. Identify and explain the key clinical features that may be associated with patient mortality outcomes through the “triangulation”, which will provide directions for subsequent mechanistic studies and larger-scale validation studies.

## Method

### Data Collection

#### Patients

This study retrospectively analyzed patients with grade 4 irAEs admitted to the ICU of Tianjin Medical University Cancer Institute & Hospital from January 2018 to December 2024. All participants received PD-1/PD-L1 inhibitor therapy prior to ICU transfer for critical care. The study was approved by the Ethics Committee of Tianjin Medical University Cancer Institute & Hospital. Due to the retrospective nature of the study, the requirement for informed patient consent was waived by the committee. All patient data were anonymized prior to analysis. Diagnosis of grade 4 irAEs as the primary etiology was independently confirmed by two ICU physicians, and disagreements were resolved by a third clinician. The grading for irAEs was defined based on the NCCN Guidelines for management of ICI toxicities. (1) ICI-pneumonia is typically identified as focal or diffuse inflammation of the lung parenchyma on CT imaging. Symptoms may include dry cough, shortness of breath, fever, chest pain, and increased oxygen requirement. The imaging features of pneumonitis are known to be variable and may include ground-glass opacities, organizing pneumonia, hypersensitivity, reticulonodular changes, or a mixture of all these manifestations. The grade 4 for immune-related pneumonia are life-threatening respiratory compromise.<sup>24</sup> (2) ICI-myocarditis has nonspecific, rare, and potentially severe symptoms. Etiologically, ICI-myocarditis is not viral. It is associated with myositis/myasthenia gravis, and is typically observed in combination therapy. In fatal cases, conduction abnormalities are the main death-related cause, with preserved ejection fraction. Grade 4, moderate to severe decompensation (worsening of signs and symptoms), hemodynamic instability (hypotension/cardiomyopathy), cardiac biomarkers (creatinine kinase and troponin levels)  $>3 \times$  upper limit of the normal range, life-threatening; urgent intervention indicated.<sup>25</sup> (3) Dermatologic toxicities are the most common irAEs that occur with ICIs, with majority being low-grade; however, some may be severe, with a debilitating effect on quality of life. Earlier versions of the guidelines include recommendations for the following dermatologic irAEs: maculopapular rash, pruritus, blistering disorders, Stevens-Johnson syndrome (SJS), and toxic epidermal necrolysis (TEN). Since only SJS and TEN are involved in this study, the grading criteria for SJS are listed. Grade 4, skin erythema and blistering/sloughing covering  $\geq 10\%$  body surface area (BSA) with associated signs (eg, erythema, purpura, epidermal detachment, mucous membrane detachment) and/or systemic symptoms and concerning associated blood work abnormalities (eg, elevated liver function testing in the setting of DRESS/DIHS).<sup>26</sup> (4) Nephritis; Grade 4, life-threatening consequences; dialysis indicated.<sup>26</sup> (5) Guillain-Barre syndrome is progressive, most commonly characterized by symmetrical muscle weakness and absent or reduced deep tendon reflexes. Extremities, facial, respiratory, and bulbar and oculomotor nerves may be involved, and dysregulation of autonomic nerves may exist, which usually starts with lower back and thigh pain. Grade 3 and Grade 4, limited self-care and aids warranted, weakness limiting waking, any dysphagia, facial weakness, respiratory muscle weakness, or rapidly progressive symptoms.<sup>24</sup> (6) Transaminitis without elevated bilirubin requires elevated alanine transaminase (ALT) and aspartate transaminase (AST); Grade 4, life-threatening  $> 20 \times$  ULN.<sup>24</sup>

## Clinical Data Collection

Based on clinical outcomes, patients were stratified into Survivors (n=13; stabilized and transferred to oncology wards for continued therapy) and non-Survivors (n=13; ICU mortality). Of the 26 patients who met the inclusion criteria during the study period, exactly 13 died (non-survivors) and 13 survived (survivors). This equal split occurred by chance and was not a predefined matching or equal-sampling design. Data collection: (1) Retrospective medical record review of ICI treatment history. (2) Structured interviews with patients/families to obtain past medical history. (3) Biochemical profiling via blood samples drawn within 15 minutes of admission to the ICU (radial artery or central venous catheterization). Patients who met the inclusion criteria were retrospectively identified from the ICU database. After identification, all enrolled patients were prospectively followed for 12 months according to a predefined protocol, and data were systematically recorded in a dedicated database. Thus, this study is a retrospective cohort study with prospective follow-up of identified cases.

## Traditional Statistical Methods Analysis

Descriptive data of continuous variables were presented as medians with Interquartile Ranges (IQR) and the frequencies of categorical variables were expressed as percentages. Intergroup comparison was performed using non-parametric methods: Mann–Whitney *U*-test for continuous variables and either Pearson's chi-square test or Fisher's exact test for categorical variables, with the latter selected when the expected cell frequencies fell below 5. All statistical analyses were conducted using IBM® SPSS® Statistics for Windows, version 25.0 (IBM Corp., Armonk, N.Y., USA). *p*-values were corrected for multiple comparisons using Fisher's Least Significant Difference (LSD) method. The LSD method controls the type I error at the comparison-wise level rather than the family-wise level, which is more appropriate for hypothesis-generating research where the goal is to avoid missing potentially relevant associations. A two-sided *p*-value of <0.05 after LSD correction was considered statistically significant.

## Machine Learning Analysis

Given the high-dimensional (n = 75) and small-sample (n = 26) nature of the data, we employed a machine learning pipeline designed to maximize robustness and minimize overfitting. The strategies like Leave-One-Out Cross-Validation (LOOCV) combined with Bootstrap resampling for confidence interval estimation can maximize the robustness of model evaluation using limited data, providing a feasible pathway for exploratory research. Specifically, the analytical methods used in this study are as follows:

For continuous variables, missing values were imputed using the mean; for non-continuous variables, missing values were imputed using the mode. Missing value imputation was performed only for the machine learning analysis described in this section. Traditional statistical analyses were conducted on the original data without any imputation. Among these, 6 features contain missing values, and the number of missing values for all features is less than 2. The initial feature pool comprises 75 features, far exceeding the sample size (n = 26). To prevent the curse of dimensionality and mitigate the risk of model overfitting, it is inadvisable to directly employ all features for modeling. In this case, LOOCV was employed to enhance the robustness of model evaluation. In this method, each iteration uses 25 samples for training and the remaining 1 sample for testing, repeated 26 times until every sample has been used once for testing. For model selection, we included six classic machine learning algorithms: Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), Naive Bayes (NB), and Linear Discriminant Analysis (LDA), to cover a range of different modeling assumptions and structural characteristics. The final prognostic outcome served as the target variable for prediction.

To comprehensively compare model performance, we selected six common evaluation metrics: Area Under the Curve (AUC), accuracy, precision, recall, specificity, and the F1-score. The mean value for each metric was calculated based on the 26 cross-validation iterations. Recognizing that standard deviations can be less robust in small-sample models, we calculated 95% confidence intervals for each performance metric using the Bootstrap resampling method with 2000 resamples. Based on the evaluation results, models with poor or unstable performance were discarded, while those demonstrating a certain level of discriminative ability were retained for subsequent analysis.

For the retained models, we retrained them using the entire dataset and performed SHAP (SHapley Additive exPlanations) analysis to obtain the corresponding feature importance ranking for each model. To comprehensively evaluate feature importance, we normalized the SHAP values from each model, then calculated their average to derive an overall ranking of feature contributions. Based on the overall feature contribution ranking, the difference in SHAP values was calculated to find the steepest descending point, and this was used to determine the number of retained features, to retain the corresponding number of features in each model versus the overall ranking, and to filter out the features that were repeated in multiple models (number of occurrences  $\geq 2$ ), constituting a subset of the features that were ultimately used for modeling and interpretation.

Finally, the selected feature subset was re-input into the retained machine learning models. The model with the best performance was chosen as the final model. Based on its SHAP results, the final feature importance ranking and model interpretation were derived to enhance model explainability and the reliability of the results.

## Triangulation

In high-dimensional, low-sample-size scenarios, traditional statistical methods often exhibit instability, while machine learning approaches are prone to overfitting. To address this, we introduce triangular verification as a strategy for feature dimensionality reduction and stability screening. This approach integrates traditional statistical methods with machine learning techniques, extracting key features identified by each and selecting their intersection as the final feature set. Furthermore, features that fail to meet traditional statistical significance criteria but are deemed interpretatively relevant through machine learning SHAP analysis are excluded, thereby enhancing the robustness and interpretability of feature selection.

## Result

### Traditional Statistical Results

Based on the final prognostic outcomes of the patients, we divided the 26 enrolled patients into a survival group and a non-survival group. After collecting basic information, clinical characteristic indicators, laboratory test results, treatment parameters, and diagnostic data from these 26 patients during their ICU stay, a database was established. Following conventional statistical analysis, clinically relevant features with statistically significant differences were identified, as shown in [Table 1](#). Statistics for other indicators that did not show significant differences are detailed in [Tables S1–S3](#). However, it should be noted that due to the very small sample size ( $n=26$ ), univariate comparisons, especially those with small cell counts, should be interpreted with extreme caution. These results are intended for hypothesis generation only and require validation in larger cohorts.

First, compared to the non-survival group, the survival group had a lower proportion of males (53.85% vs. 92.31%) and a higher body mass index (BMI) ( $25.06 \pm 4.17$  vs.  $22.05 \pm 2.20$ ). Regarding tumor pathological type, the proportion of squamous cell carcinoma was higher in the survival group than in the non-survival group (61.54% vs. 30.77%). In terms of medical oncology treatment, compared to the non-survival group, the survival group had a higher proportion of patients who had received vascular endothelial growth factor (VEGF) inhibitors (53.85% vs. 15.38%), a higher proportion without a history of surgery (76.92% vs. 38.46%), and a lower proportion in the perioperative period (15.38% vs. 53.85%).

In this study, a relatively high proportion of deaths in the non-survival group were related to adverse events occurring in the lungs, accounting for 69.23%, whereas the incidence of pneumonia in the survival group was only 7.69%. This suggests an extremely high mortality rate among patients with grade 4 adverse event pneumonia. In contrast, the proportion of patients transferred to the ICU primarily due to acute respiratory failure was significantly higher in the survival group than in the non-survival group (76.92% vs. 30.77%). This may indicate that intervention for respiratory failure in the ICU can improve the prognosis of patients with grade 4 adverse events, whereas mortality rises sharply once severe pneumonia develops.

Regarding clinical indicators, the survival group had significantly higher diastolic blood pressure, arterial blood pH, and lymphocyte count compared to the non-survival group. In contrast, the neutrophil percentage, lactate levels over the first three consecutive days after ICU admission, as well as total bilirubin and direct bilirubin levels, were significantly lower in the survival group.

**Table 1** Characteristics of Critically Ill Cancer Patients with Grade 4 irAEs

	All irAEs (n=26)	Survivors (n=13)	Non-Survivors (n=13)	p
Sex=M (n, %)	19(73.1%)	7(53.85%)	12(92.31%)	0.027
Sex=F (n, %)	7(26.9%)	6(46.15%)	1(7.69%)	0.027
BMI (kg/m <sup>2</sup> , $\bar{x} \pm s$ )	23.56±3.61	25.06±4.17	22.05±2.20	0.030
VEGF-inhibitors (n, %)	9(34.62%)	7(53.85%)	2(15.38%)	0.039
Squamous carcinoma (n, %)	12(46.15%)	8(61.54%)	4(30.77%)	0.047
No surgical experience (n, %)	15(57.7%)	10(76.92%)	5(38.46%)	0.047
Postoperative (n, %)	9(34.6%)	2(15.38%)	7(53.85%)	0.039
Pneumonia (n, %)	10(38.46%)	1(7.69%)	9(69.23%)	0.001
Acute respiratory failure to ICU (n, %)	14(53.85%)	10(76.92%)	4(30.77%)	0.014
Other reason to ICU (n, %)	12(46.15%)	2(15.38%)	10(76.92%)	0.002
DBP [mmHg, M (Q1, Q3)]	75.5[60–82.25]	79[72.5–87]	62[58–77]	0.014
pH [M (Q1, Q3)]	7.36[7.34–7.46]	7.41[7.31–7.48]	7.35[7.33–7.41]	0.029
N% [%; M (Q1, Q3)]	89.35[82.58–94.78]	84.9[75.5–90.65]	93.8[89.3–96.15]	0.009
LAC at d <sub>1</sub> [mmol/L, M (Q1, Q3)]	2[1.68–4.23]	1.7[1.05–2.2]	2.5[1.9–5.35]	0.006
LAC at d <sub>2</sub> [mmol/L, M (Q1, Q3)]	1.8[1.48–2.48]	1.5[1.25–1.95]	2[1.8–4.2]	0.007
LAC at d <sub>3</sub> [mmol/L, M (Q1, Q3)]	1.6[1.28–2.1]	1.5[1.1–1.75]	1.9[1.45–4.2]	0.026
TBIL [ $\mu$ mol/L, M (Q1, Q3)]	10.6[7.88–15.88]	8.4[6.8–10.6]	13.7[10.4–22.55]	0.005
DBIL [ $\mu$ mol/L, M (Q1, Q3)]	6.4[3.72–11.23]	5.1[3.15–6.4]	9.6[5.65–13.15]	0.039
Nasal catheter oxygen inhalation (n, %)	8(30.8%)	6(46.15%)	2(15.38%)	0.005
Mechanical ventilation (n, %)	14(53.8%)	3(23.08%)	11(84.62%)	0.002
FiO <sub>2</sub> [%; M (Q1, Q3)]	40[29–100]	29[29–40]	100[40–100]	0.002
Vasopressor therapy (n, %)	17(65.4%)	4(38.46%)	13(100.00%)	0.000

**Abbreviations:** BMI, Body mass index; VEGF-inhibitors, Vascular endothelial growth factor; DBP, Diastolic blood pressure; N%, Percentage of neutrophils; LAC, Lactic acid; TBIL, Total bilirubin; DBIL, Direct bilirubin; FiO<sub>2</sub> Fraction of inspiration O<sub>2</sub>.

In terms of ICU management, the study showed that the non-survival group required significantly more advanced oxygen therapy modalities—such as high-flow humidified oxygen therapy, ventilator-assisted ventilation, and higher inspired oxygen concentrations—than the survival group, demonstrating greater oxygen dependency. Furthermore, the proportion of patients receiving vasoactive drugs for blood pressure support was lower in the survival group compared to the non-survival group (38.46% vs. 100.00%).

## Model Screening and Performance Comparison

To evaluate model performance under limited sample sizes, we constructed models on the full feature set and tested their performance on the training set. We systematically assessed six machine learning models using leave-one-out cross-validation. The predictive performance of each model is detailed in [Table 2](#) and [Figure 1](#).

Using the full feature set to build the model resulted in significant overfitting. Among them, the DT, NB, and LDA models were excluded as their AUC confidence intervals included 0.5 and their overall performance was poor. Consequently, the SVM, RF, and LR models were retained for subsequent feature analysis.

The SVM, RF, and LR models were trained using the entire sample set (n=26). Feature ranking was then performed through SHAP analysis. A significant steepening point ( $\Delta = 0.045$ ) was found at the 11<sup>th</sup> in the ranking. Therefore, the top 11 features were retained. Among them, each of the top 11 features ranked by the mean SHAP value appeared at least once within the top 11 features of the other three models, while no other features overlapped. Therefore, the following 11 features were ultimately selected as the final feature subset for modeling and interpretation: Oxygen therapy, BMI, LAC1, Vasopressor therapy, LAC2, VEGF-inhibitors, Uric acid, N%, Targeted therapy, Age and Pneumonia. The complete ranking of all features is detailed in [Table S4](#).

**Table 2** The Prediction Performance of Six Models

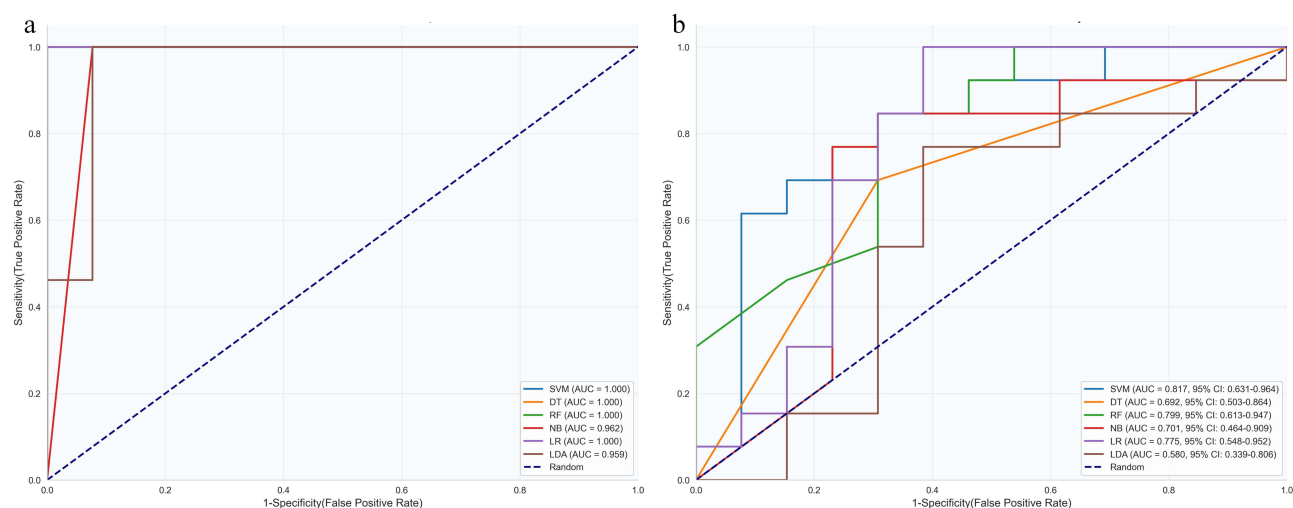
Model	AUC	Accuracy	Precision	Recall	Specificity	FI
<b>SVM</b>						
Train	1.00	1.00	1.00	1.00	1.00	1.00
LOOCV	0.82(0.63–0.96)	0.77(0.62–0.92)	0.73(0.50–0.93)	0.85(0.64–1.00)	0.69(0.42–0.92)	0.79(0.59–0.93)
<b>DT</b>						
Train	1.00	1.00	1.00	1.00	1.00	1.00
LOOCV	0.69(0.50–0.87)	0.69(0.50–0.85)	0.69(0.43–0.93)	0.69(0.42–0.93)	0.69(0.42–0.93)	0.69(0.45–0.87)
<b>RF</b>						
Train	1.00	1.00	1.00	1.00	1.00	1.00
LOOCV	0.80(0.60–0.95)	0.73(0.54–0.88)	0.69(0.46–0.89)	0.84(0.62–1.00)	0.62(0.33–0.87)	0.76(0.55–0.90)
<b>NB</b>						
Train	0.96	0.92	0.93	0.92	0.85	0.92
LOOCV	0.70(0.47–0.92)	0.73(0.54–0.88)	0.75(0.45–1.00)	0.69(0.43–0.93)	0.77(0.50–1.00)	0.72(0.47–0.89)
<b>LR</b>						
Train	1.00	1.00	1.00	1.00	1.00	1.00
LOOCV	0.77(0.57–0.94)	0.73(0.54–0.88)	0.71(0.44–0.93)	0.78(0.50–1.00)	0.69(0.43–0.92)	0.74(0.52–0.90)
<b>LDA</b>						
Train	0.96	0.92	0.92	0.92	0.92	0.92
LOOCV	0.58(0.35–0.82)	0.58(0.38–0.77)	0.54(0.25–0.81)	0.54(0.27–0.82)	0.62(0.33–0.87)	0.56(0.29–0.76)

**Abbreviations:** LOOCV, leave-one-out cross-validation; CI, confidence interval; SVM, support vector machine; RF, random forest; LR, logistic regression; DT, decision tree; NB, naive Bayes; LDA, linear discriminant analysis; AUC, Area Under the Curve; FI, F1 Score.

## Final Model Construction and Interpretation

Three machine learning models—SVM, RF, and LR—were employed to construct models using the final set of 11 selected features. The construction method followed the procedure described in Model Screening and Performance Comparison, where performance metrics were calculated via leave-one-out cross-validation to determine the model for interpretation. The models were built using the entire sample set and were interpreted via SHAP analysis. The predictive performance of the three models and their corresponding confidence intervals are presented in Table 3 and Figure 2.

For the 11 selected features, all models demonstrated satisfactory classification performance. Based on a comprehensive evaluation of all metrics, the LR model demonstrated the best performance. The LR model was retrained using the entire sample set and subjected to SHAP analysis, with the final results illustrated in Figure 3.



**Figure 1** ROC curves for the performance of six models. (a) ROC of the train subset. (b) ROC of the LOOCV.

**Abbreviations:** LOOCV, leave-one-out cross-validation; CI, confidence interval; SVM, support vector machine; RF, random forest; LR, logistic regression; DT, decision tree; NB, naive Bayes; LDA, linear discriminant analysis.

**Table 3** The Prediction Performance and Confidence Intervals of SVM-RF-LR Models

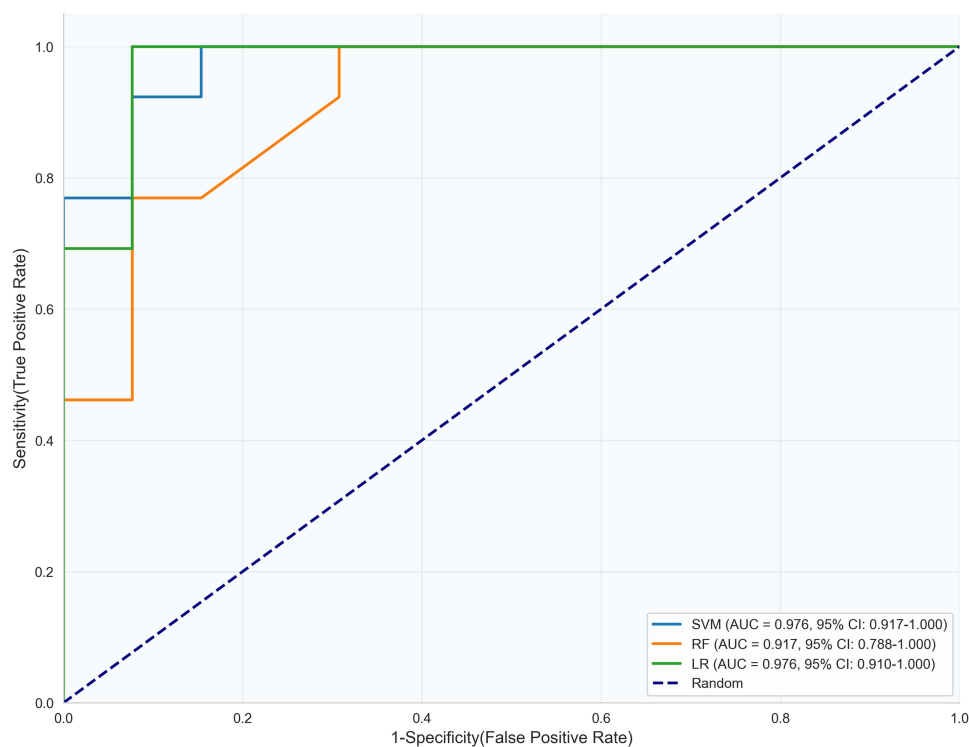
Model	AUC (95% CI)	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	Specificity (95% CI)	FI (95% CI)
SVM	0.98 (0.91–1.00)	0.84 (0.69–0.96)	0.91 (0.70–1.00)	0.77 (0.53–1.00)	0.92 (0.75–1.00)	0.83 (0.64–0.96)
RF	0.92 (0.79–1.00)	0.81 (0.65–0.96)	0.79 (0.53–1.00)	0.85 (0.62–1.00)	0.77 (0.50–1.00)	0.81 (0.61–0.96)
LR	0.98 (0.91–1.00)	0.88 (0.73–1.00)	0.92 (0.73–1.00)	0.85 (0.63–1.00)	0.92 (0.75–1.00)	0.88 (0.71–1.00)

**Abbreviations:** CI, confidence interval; SVM, support vector machine; RF, random forest; LR, logistic regression; AUC, Area Under the Curve; FI, FI Score.

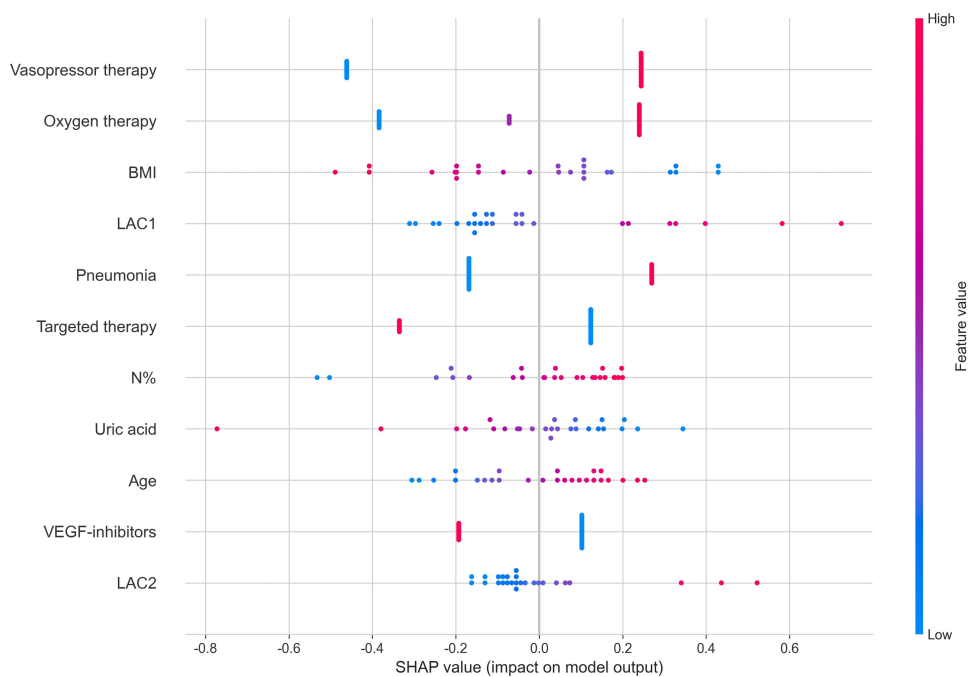
It can be observed that each feature contributed to a certain extent to the final model. Among them, BMI, Targeted therapy, Uric acid and VEGF-inhibitors showed a negative correlation with the outcome variable, while Vasopressor therapy, Oxygen therapy, LAC1, Pneumonia, N%, Age, and LAC2 exhibited a positive correlation with the outcome variable. Among these, Oxygen therapy was a particularly unique approach. Nasal catheter oxygen inhalation and high flow rate humidification were associated with favorable outcomes, while mechanical ventilation typically predicted adverse outcomes.

## Triangulation

A total of 21 features were found to be significantly different ( $p < 0.05$ ) in traditional statistical methods. In machine learning, a total of 11 features were retained. In order to unify and integrate the features, the features of traditional statistical methods were grouped and merged. Among them, Squamous carcinoma was grouped into Pathological state; No surgical experience and Postoperative were merged into Surgical operation condition; Acute respiratory failure to ICU and Other reason to ICU were combined as Reason for being transferred to the ICU; Nasal catheter oxygen inhalation and Mechanical ventilation were combined as Oxygen therapy. Finally, 18 characteristics were obtained. There was significant overlap in the set of features identified by the two independent analysis methods, which together formed a final feature set of 8 features, as shown in Figure 4. The correlation between the overlapping 8 features and the final death outcome was consistent across the two methods, making these 8 features important for the final outcome. Among

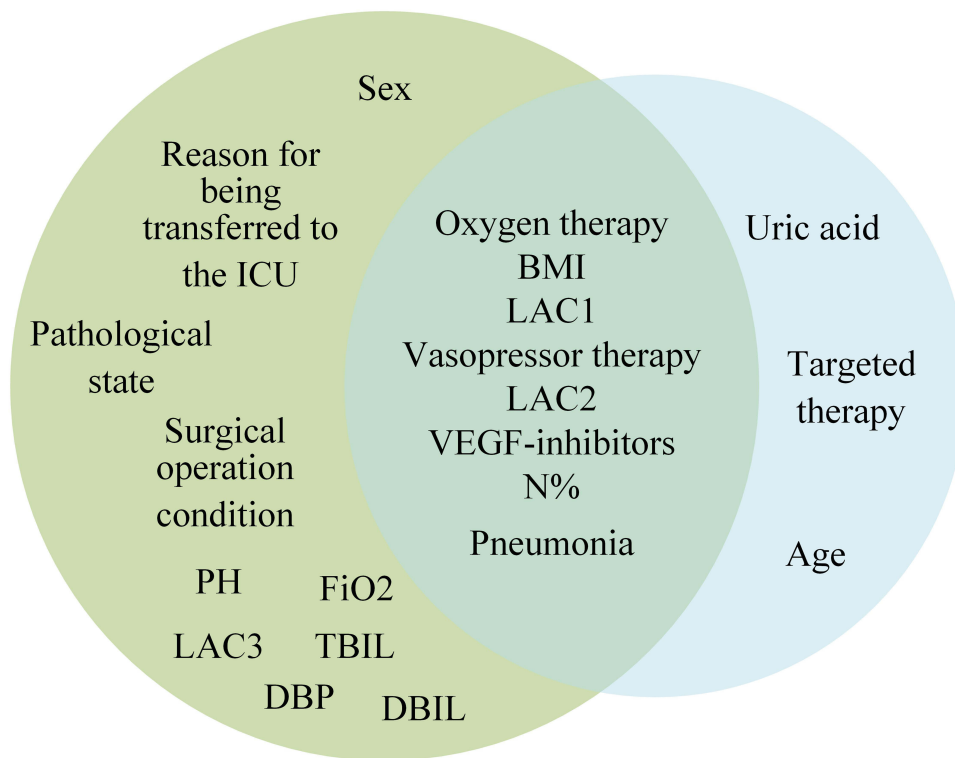
**Figure 2** ROC curves for the performance of SVM-RF-LR Models.

**Abbreviations:** CI, confidence interval; SVM, support vector machine; RF, random forest; LR, logistic regression.



**Figure 3** SHAP explanation plot.

**Abbreviations:** BMI, Body mass index; LAC1, Lactic Acid at d1; LAC2, Lactic Acid at d2; N%, Percentage of neutrophils; VEGF-inhibitors, Vascular endothelial growth factor inhibitors.



**Figure 4** Venn diagram of Triangulation.

**Abbreviations:** BMI, Body mass index; VEGF-inhibitors, Vascular endothelial growth factor; DBP, Diastolic blood pressure; N%, Percentage of neutrophils; LAC, Lactic acid; TBIL, Total bilirubin; DBIL, Direct bilirubin; FiO2 Fraction of inspiration O2.

them, BMI and VEGF-inhibitors showed a negative correlation with the outcome variable, while Vasopressor therapy, Oxygen therapy, LAC1, Pneumonia, N%, and LAC2 exhibited a positive correlation with the non-Survivors outcome. Notably, none of the 6 features that had missing values entered this final 8-feature set, further mitigating concerns about imputation-induced bias.

## Discussion

This exploratory preliminary study targeted the clinical population of Grade 4 irAEs, which was high-risk but understudied. Faced with the practical challenge of extremely small sample size, this study did not attempt to construct a clinically predictive model with generalizability. Instead, a more fundamental question, which clinical characteristics may be robustly associated with outcomes among critically ill patients admitted to the ICU due to Grade 4 irAEs, was investigated. To this end, we innovatively applied the research strategy of “triangulation”, integrating traditional statistical methods with machine learning techniques to maximize the extraction of reliable information from limited data, including 26 patients admitted to the ICU between 2018 and 2024. Ultimately, we identified 8 dual-validated high-confidence candidate features: vasopressor therapy, oxygen therapy, body mass index (BMI), lactate acid level at Day1 and Day2 (LAC1, LAC2), neutrophil percentage (N%), history of VEGF-inhibitor use, and pneumonia (as an irAEs type). But it must be emphasized that this is an exploratory study, all findings are preliminary and require replication.

Our findings both align with and extend the current understanding of severe irAEs. Consistent with prior research, our conventional statistical analysis confirmed that severe irAEs involving the lungs (specifically, grade 4 pneumonia) are associated with an alarmingly high mortality rate, which aligns with reports of fatal toxic effects being strongly linked to specific organ systems like pneumonitis<sup>27</sup> and myocarditis.<sup>28</sup> Furthermore, the association between features like elevated lactate levels, neutrophilia, the need for vasopressors and high levels of respiratory support with poor outcomes echoes established prognostic indicators in critical care and sepsis, underscoring that the final common pathway of fatal irAEs often converges on multi-organ failure and shock.<sup>29</sup> The importance of serial lactate measurements (LAC1 and LAC2) highlights that not just the initial value, but the trend and response to early resuscitation are critical.

However, our study yields nuanced insights that diverge from some broader analyses of irAEs. While several studies combining all-grade irAEs suggest that their occurrence may correlate with better antitumor response,<sup>30,31</sup> our focused analysis on grade 4 events reveal a starkly different clinical reality dominated by life-threatening complications. This discrepancy highlights the critical importance of grading severity and studying high-grade events separately, as pooling data can mask the distinct risk profiles and outcomes of the most severe cases. Additionally, this study prioritized a combination of features (eg, BMI, VEGF-inhibitors) not commonly highlighted as top predictors in studies of lower-grade or mixed-severity irAEs. This suggests that the pathophysiology and prognostic drivers of grade 4 events may be distinct. Higher BMI demonstrated a positive correlation with survival outcomes. This aligns with certain observations of the “obesity paradox”<sup>32</sup> in critically ill patients in general, potentially suggesting that greater metabolic reserves may be associated with better tolerance of severe immune stress. However, we cannot rule out the possibility that higher BMI simply reflects better nutritional reserve or less advanced disease, rather than a direct biological protective effect. Also, the use of VEGF-inhibitors was inversely associated with mortality in this cohort. The underlying mechanisms warrant further investigation. Beyond their anti-angiogenic effects, VEGF pathway inhibitors have also been reported to modulate the immune microenvironment.<sup>33</sup> In the context of severe immune toxicity, this immunoregulatory action may help mitigate excessive inflammatory responses, thereby producing a protective effect. Also, the observed survival advantage might be largely attributable to differences in tumor type rather than a direct biological protective effect of VEGF inhibitors. But due to the small sample size, we could not adequately adjust for this factor in multivariate analyses. Of course, this association may also be influenced by unmeasured confounding factors, necessitating validation through larger prospective studies in the future.

It is worth noting that while not filtered through the most stringent “triangular verification” process, other features identified solely through traditional statistics or machine learning still hold significant indicative value. These may reflect associations with mortality outcomes across different dimensions or fail to demonstrate consistent robustness due to sample size limitations. The other 10 features shown significant differences in traditional statistics suggested non-survivors may exhibit more aggressive tumor biology, such as male gender<sup>34</sup> and postoperative. Also, they may exhibit more severe acidosis (lower pH), poorer physiological function status (lower DBP, higher FiO2 and Acute respiratory failure to ICU) and greater hepatic injury (higher

TBIL and DBIL). In the result of machine learning, uric acid, targeted therapy and age were identified as significant features. Although they did not reach statistical significance in traditional analyses, they may influence outcomes through complex interactions with other characteristics, particularly age as a fundamental physiological indicator. These “secondary” feature sets not only deepen our understanding of the clinical characteristics of critically ill patients with Grade 4 immune-related adverse events but also provide a source of testable hypotheses for subsequent large-scale studies. Furthermore, the discrepancies in results across different methods precisely highlight the advantage of multi-method integration—its ability to capture potential biological signals across diverse analytical dimensions. This finding warrants focused exploration in future research. Therefore, it is necessary to further validate and explore these discoveries by expanding the sample size.

Compared with previous studies, our feature identification shows some differences. Acar found that baseline eosinophilia and low RDW predicted the risk of grade  $\geq 3$  irAEs occurrence.<sup>20</sup> Due to data collection limitations, RDW was not captured in our study, and eosinophils were not retained in the final 8-feature set. This discrepancy likely stems from differences in outcomes and populations: Acar et al focused on the incidence of irAEs, whereas we examined mortality in grade 4 irAEs; moreover, all our patients were critically ill requiring ICU care, likely in a decompensated immune state where early predictors are overshadowed by more direct indicators of organ failure. These comparisons suggest that predicting “risk of irAEs occurrence” versus “risk of death in grade 4 irAEs” may require different feature sets, highlighting the necessity of studying this extreme subgroup independently.

Although this study is exploratory in nature, our findings still hold clear clinical significance. First, all 8 high-confidence features identified are routine indicators in daily ICU monitoring. Clinicians should remain highly vigilant for such feature combinations—particularly when vasopressor requirements, elevated lactate levels, and high respiratory support levels coexist—as they serve as strong indicators of critical illness and extremely high mortality risk in Grade 4 irAEs patients. Also, findings reinforce that early, aggressive hemodynamic support (correcting hypotension, lowering lactate) and respiratory support from the cornerstone of management for Grade 4 irAEs patients. Concurrently, patients receiving VEGF inhibitors warrant closer monitoring. What’s more, this study not only provides direction for subsequent validation of the core 8 features but also paves the way for exploring the relationship between features suggested by single methods (such as TBIL, uric acid, specific treatment history) and prognosis, as well as their interactions with core features.

This study has several limitations that must be acknowledged. First, the small sample size ( $n = 26$ ), while inevitable given the rarity of grade 4 irAEs requiring ICU care, limits the statistical power of conventional analyses and increases the risk of overfitting in machine learning models. We mitigated this through LOOCV, Bootstrap confidence intervals and “triangulation”, but the results require external validation in larger, multi-center cohorts. Also, although the proportion of missing data was very low, simple imputation in a small sample may still introduce uncertainty. Future studies are encouraged to apply more advanced methods. However, the triangulation framework ensured that only features consistently significant across both statistical and machine learning pathways were retained, and all features with missing values were excluded from the final feature set. This design reduces, to some extent, the risk that imputation-related artifacts drove our conclusions. Second, the retrospective, single-center design may introduce selection bias and limit the generalizability of our findings to other patient populations or healthcare settings. Third, our model is based on data available at or shortly after ICU admission. Incorporating dynamic data trends over the first 24–72 hours of ICU stay could potentially improve predictive accuracy. And our study did not include a control group of ICU patients without irAEs. Therefore, we cannot distinguish between irAEs-specific prognostic factors and general markers of critical illness severity. Future studies should include such a control group to address this question.

In spite of these limitations, this study offers a pragmatic and important step forward. It demonstrates the feasibility and potential utility of applying interpretable machine learning to “small data” challenges in oncology-critical care. The final 8 features are variables readily available in any ICU, making it potentially translatable to bedside risk assessment. Clinically, our findings reinforce the paramount importance of early recognition and aggressive supportive care for cardiovascular and respiratory instability in grade 4 irAEs. This lays the groundwork for developing tools to identify patients at highest risk of death upon ICU admission, enabling more intensive monitoring and prompt consideration of second-line immunomodulatory therapies.

Future studies should focus on multi-center collaborations to assemble larger datasets for refining and validating this model. Moreover, data collection should be substantially improved in terms of granularity, including systematic registration of

primary tumor type, treatment setting, pathological subtype, complete histories of concomitant chemotherapy and so on. Only with such detailed information can the independent predictive value of the features identified in this exploratory study be properly assessed. Prospective studies are needed to evaluate the impact of using such a predictive tool on clinical decision-making and patient outcomes.

## Conclusion

In conclusion, this explicitly exploratory study employed a robustness-based discovery framework utilizing “triangulation” to preliminarily identify a high-confidence candidate feature set comprising 8 clinical characteristics closely associated with mortality outcomes in the high-risk population of Grade 4 irAEs. It also systematically examined the indicative significance of other potentially associated features. But these findings are strictly hypothesis-generating and must not be interpreted as practice-changing. By focusing on this extreme-risk subgroup and employing robust analytical methods tailored for small samples, we identified a parsimonious set of clinical predictors. Our findings highlight the distinct and perilous nature of grade 4 irAEs, characterized by profound shock and respiratory failure, and underscore the critical need for specialized prognostic tools and management protocols for this vulnerable population. While all identified features should be considered hypothesis-generating and require rigorous external validation in larger independent cohorts, this work provides a foundational framework for risk stratification and motivates further investigation into the biology and management of life-threatening irAEs.

## Data Sharing Statement

The datasets used and analyzed during the current study are available from Professor Yanmei Zhao (corresponding author) upon reasonable request.

## Ethics Approval and Consent to Participate

The data collection and processing protocols were performed in accordance with the Declaration of Helsinki, approved by the institutional ethics committee (Ethics Committee of Tianjin Medical University Cancer Institute & Hospital, bc20251790). The following work was carried out before the official launch of this study: 1. Registration was conducted through the ethics committee. 2. Exemption from informed consent was granted.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This work was supported by the Project of integrated traditional Chinese and Western Medicine of Tianjin Health Commission (2023157, 2023171), Funded by Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-009A), Scientific research topic of TCM in Hebei Province Administration (T2025061) and Tianjin Municipal Commission of Science and Education Research Program General Project (2025KJ074).

## Disclosure

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

1. Tan S, Day D, Nicholls SJ, Segelov E. Immune checkpoint inhibitor therapy in oncology: current uses and future directions: JACC: cardioOncology State-of-the-Art Review. *JACC CardioOncol.* 2022;4(5):579–597. doi:10.1016/j.jacc.2022.09.004
2. Ledford H, Else H, Warren M. Cancer immunologists scoop medicine Nobel prize. *Nature.* 2018;562(7725):20–21. doi:10.1038/d41586-018-06751-0

3. Reck M, Rodriguez-Abreu D, Robinson AG, et al. Updated analysis of KEYNOTE-024: pembrolizumab versus platinum-based chemotherapy for advanced non-small-cell lung cancer with PD-L1 tumor proportion score of 50% or greater. *J Clin Oncol.* 2019;37(7):537–546. doi:10.1200/JCO.18.00149
4. Choueiri TK, Motzer RJ. Systemic therapy for metastatic renal-cell carcinoma. *New Engl J Med.* 2017;376(4):354–366. doi:10.1056/NEJMra1601333
5. Ascierto PA, Del Vecchio M, Mandalá M, et al. Adjuvant nivolumab versus ipilimumab in resected stage IIIB-C and stage IV melanoma (CheckMate 238): 4-year results from a multicentre, double-blind, randomised, controlled, Phase 3 trial. *Lancet Oncol.* 2020;21(11):1465–1477. doi:10.1016/S1470-2045(20)30494-0
6. Haslam A, Prasad V. Estimation of the percentage of US patients with cancer who are eligible for and respond to checkpoint inhibitor immunotherapy drugs. *JAMA Netw Open.* 2019;2(5):e192535. doi:10.1001/jamanetworkopen.2019.2535
7. Alsaab HO, Sau S, Alzhrani R, et al. PD-1 and PD-L1 checkpoint signaling inhibition for cancer immunotherapy: mechanism, combinations, and clinical outcome. *Front Pharmacol.* 2017;8:561. doi:10.3389/fphar.2017.00561
8. von Itzstein MS, Gerber DE, Bermas BL, Meara A. Acknowledging and addressing real-world challenges to treating immune-related adverse events. *J Immuno Ther Cancer.* 2024;12(7):e009540. doi:10.1136/jitc-2024-009540
9. von Itzstein MS, Khan S, Gerber DE. Investigational biomarkers for checkpoint inhibitor immune-related adverse event prediction and diagnosis. *Clin Chem.* 2020;66(6):779–793. doi:10.1093/clinchem/hvaa081
10. Postow MA, Sidlow R, Hellmann MD. Immune-related adverse events associated with immune checkpoint blockade. *New Engl J Med.* 2018;378(2):158–168. doi:10.1056/NEJMra1703481
11. Song P, Zhang D, Cui X, Zhang L. Meta-analysis of immune-related adverse events of immune checkpoint inhibitor therapy in cancer patients. *Thoracic Cancer.* 2020;11(9):2406–2430. doi:10.1111/1759-7714.13541
12. Wang Y, Zhou S, Yang F, et al. Treatment-related adverse events of PD-1 and PD-L1 inhibitors in clinical trials: a systematic review and meta-analysis. *JAMA Oncol.* 2019;5(7):1008–1019. doi:10.1001/jamaoncol.2019.0393
13. Sznol M, Ferrucci PF, Hogg D, et al. Pooled analysis safety profile of nivolumab and ipilimumab combination therapy in patients with advanced melanoma. *J Clin Oncol.* 2017;35(34):3815–3822. doi:10.1200/jco.2016.72.1167
14. Tamura K, Takenaka Y, Hosokawa K, et al. Effect of immortal time bias on the association between immune-related adverse events and oncological outcomes following immune checkpoint inhibitors therapy for head and neck squamous cell carcinoma. *PLoS One.* 2024;19(11):e0314209. doi:10.1371/journal.pone.0314209
15. Talab TJ, Shafi FAA, Jasim SM. Association between immune-related adverse events and therapeutic outcomes in patients with advanced non-small cell lung cancer treated with pembrolizumab: a retrospective study. *World Acad Sci J.* 2025;7(5):82. doi:10.3892/wasj.2025.370
16. Wang DY, Salem JE, Cohen JV. Fatal toxic effects associated with immune checkpoint inhibitors: a systematic review and meta-analysis (vol 4, pg 1721, 2018). *JAMA Oncol.* 2018;4(12):1792. doi:10.1001/jamaoncol.2018.5346
17. Wang DY, Salem JE, Cohen JV, et al. Fatal toxic effects associated with immune checkpoint inhibitors: a systematic review and meta-analysis. *JAMA Oncol.* 2018;4(12):1721–1728. doi:10.1001/jamaoncol.2018.3923
18. Haanen JBAG, Carbone F, Robert C, et al. Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2017;28:119–142. doi:10.1093/annonc/mdx225
19. Guezour N, Soussi G, Brosseau S, et al. Grade 3–4 immune-related adverse events induced by immune checkpoint inhibitors in Non-Small-Cell Lung Cancer (NSCLC) patients are correlated with better outcome: a real-life observational study. *Cancers.* 2022;14(16):3878. doi:10.3390/cancers14163878
20. Acar C, Açar FP, Şahin G, Yüksel HÇ, Karaca B, Göker E. Using routine blood tests to predict severe immune-related adverse events during immune checkpoint inhibitor treatment. *BMC Cancer.* 2025;26(1):83. doi:10.1186/s12885-025-15460-7
21. De Velasco G, Je YJ, Bossé D, et al. Comprehensive meta-analysis of key immune-related adverse events from CTLA-4 and PD-1/PD-L1 inhibitors in cancer patients. *Cancer Immunol Res.* 2017;5(4):312–318. doi:10.1158/2326-6066.Cir-16-0237
22. Cavalheiro LP, Bernard S, Barddal JP, Heutte L. Random forest kernel for high-dimension low sample size classification. *Stat Comput.* 2024;34(1). doi:10.1007/s11222-023-10309-0
23. Arshi B, Wynants L, Rijnhart E, Reeve K, Cowley LE, Smits LJ. Number of publications on new clinical prediction models: a bibliometric review. *Jmir Med Inf.* 2025;13:e62710. doi:10.2196/62710
24. Thompson JA, Schneider BJ, Brahmer J, et al. Management of immunotherapy-related toxicities. Version 1.2019. *J Natl Compr Canc Ne.* 2019;17(3):255–288. doi:10.6004/jncn.2019.0013
25. Thompson JA, Schneider BJ, Brahmer J, et al. Management of immunotherapy-related toxicities, version 1.2020. *J Natl Compr Canc Ne.* 2020;18(3):230–241. doi:10.6004/jncn.2020.0012
26. Brahmer JR, Lacchetti C, Schneider BJ, et al. Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol.* 2018;36(17):1714. doi:10.1200/Jco.2017.77.6385
27. Chen Y, Xu L, Zou S, Chen J, Xu X. Risk factors and mechanisms of immune checkpoint inhibitor-related pneumonitis. *Hum Vaccin Immunother.* 2025;21(1):2564554. doi:10.1080/21645515.2025.2564554
28. De Perna ML, Rigamonti E, Zannoni R, Espeli V, Moschovitis G. Immune checkpoint inhibitors and cardiovascular adverse events. *ESC Heart Fail.* 2025;12(4):2404–2416. doi:10.1002/ehf2.15281
29. Dijk BV, Janssen JC, Daele PLAV, et al. From ICI to ICU: a systematic review of patients with solid tumors who are treated with immune checkpoint inhibitors (ICI) and admitted to the intensive care unit (ICU). *Cancer Treat Rev.* 2025;136:102936.
30. Cook S, Samuel V, Meyers DE, et al. Immune-related adverse events and survival among patients with metastatic NSCLC treated with immune checkpoint inhibitors. *JAMA Netw Open.* 2024;7(1):e2352302. doi:10.1001/jamanetworkopen.2023.52302
31. Horvat TZ, Adel NG, Dang TO, et al. Immune-related adverse events, need for systemic immunosuppression, and effects on survival and time to treatment failure in patients with melanoma treated with ipilimumab at memorial sloan kettering cancer center. *J Clin Oncol.* 2015;33(28):3193–3198.
32. Yu L, Yuan J, Meng M, et al. Unravelling the obesity paradox in cancer: an umbrella review of protective associations and evidence credibility across 13 malignancies. *Metabolism.* 2025;156461. doi:10.1016/j.metabol.2025.156461
33. Yi M, Jiao D, Qin S, Chu Q, Wu K, Li A. Synergistic effect of immune checkpoint blockade and anti-angiogenesis in cancer treatment. *Mol Cancer.* 2019;18(1):60. doi:10.1186/s12943-019-0974-6
34. Nasca V, Zhao J, Ros J, et al. Sex and outcomes of patients with microsatellite instability-high and BRAF V600E mutated metastatic colorectal cancer receiving immune checkpoint inhibitors. *J Immuno Ther Cancer.* 2025;13(2). doi:10.1136/jitc-2024-010598

**ImmunoTargets and Therapy**

**Dovepress**  
Taylor & Francis Group

### **Publish your work in this journal**

ImmunoTargets and Therapy is an international, peer-reviewed open access journal focusing on the immunological basis of diseases, potential targets for immune based therapy and treatment protocols employed to improve patient management. Basic immunology and physiology of the immune system in health, and disease will be also covered. In addition, the journal will focus on the impact of management programs and new therapeutic agents and protocols on patient perspectives such as quality of life, adherence and satisfaction. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/immunotargets-and-therapy-journal>