Journal of Translational Medicine

RESEARCH

Open Access

Establishment and validation of predictive model of ARDS in critically ill patients

Senhao Wei[1†], Hua Zhang[2†], Hao Li[3†], Chao Li[1†], Ziyuan Shen[1], Yiyuan Yin[1], Zhukai Cong[1], Zhaojin Zeng[1], Qinggang Ge[1*], Dongfeng Li[3*] and Xi Zhu[1*]

Abstract

Background Acute respiratory distress syndrome (ARDS) is a prevalent complication among critically ill patients, constituting around 10% of intensive care unit (ICU) admissions and mortality rates ranging from 35 to 46%. Hence, early recognition and prediction of ARDS are crucial for the timely administration of targeted treatment. However, ARDS is frequently underdiagnosed or delayed, and its heterogeneity diminishes the clinical utility of ARDS biomarkers. This study aimed to observe the incidence of ARDS among high-risk patients and develop and validate an ARDS prediction model using machine learning (ML) techniques based on clinical parameters.

Methods This prospective cohort study in China was conducted on critically ill patients to derivate and validate the prediction model. The derivation cohort, consisting of 400 patients admitted to the ICU of the Peking University Third Hospital(PUTH) between December 2020 and August 2023, was separated for training and internal validation, and an external data set of 160 patients at the FU YANG People's Hospital from August 2022 to August 2023 was employed for external validation. Least absolute shrinkage and selection operator (LASSO) and multivariate logistic regression were used to screen predictor variables. Multiple ML classification models were integrated to analyze and identify the best models. Several evaluation indexes were used to compare the model performance, including the area under the receiver-operating-characteristic curve (AUC) and decision curve analysis (DCA). SHapley Additive ex Planations (SHAP) is used to interpret ML models.

Results 400 critically ill patients were included in the analysis, with 117 developing ARDS during follow-up. The final model included gender, Lung Injury Prediction Score (LIPS), Hepatic Disease, Shock, and combined Lung Contusion. Based on the AUC and DCA in the validation group, the logistic model demonstrated excellent performance, achieving an AUC of 0.836 (95% CI: 0.762–0.910). For external validation, comprising 160 patients, 44 of whom developed ARDS, the AUC was 0.799 (95% CI: 0.723–0.875), significantly outperforming the LIPS score alone.

Conclusion Combining the LIPS score with other clinical parameters in a logistic regression model provides a more accurate, clinically applicable, and user-friendly ARDS prediction tool than the LIPS score alone.

Keywords ARDS, Lung injury prediction score, Prediction model, Intensive care unit

†Senhao Wei, Hua Zhang, Hao Li and Chao Li contributed equally to this work.

*Correspondence:
Qinggang Ge
qingganggelin@126.com
Dongfeng Li
fylidongfeng@163.com
Xi Zhu
xizhuccm@163.com
Full list of author information is available at the end of the article

Background

ARDS is a syndrome characterized by acute respiratory failure resulting from various intrapulmonary or extrapulmonary factors, leading to increased permeability of the pulmonary vasculature and epithelium, pulmonary edema, gravity-dependent pulmonary atelectasis, and ultimately diffuse pulmonary inflammation and edema [1]. The LUNG SAFE study reported an incidence of ARDS in ICU patients at 10.4%, with hospital mortality rates of 34.9%, 40.3%, and 46.1% for mild, moderate, and severe ARDS [2]. A CHARDSnet study conducted in 18 ICUs across mainland China found ARDS accounted for 3.57% of total ICU admissions, with hospital mortality rates of 31.4%, 40.4%, and 56.2% for mild, moderate, and severe ARDS, respectively [3]. Current treatments primarily involve comprehensive supportive measures such as mechanical ventilation, with no drugs recognized as effective in treating ARDS, contributing to its high mortality rate. Thus, the urgent clinical need lies in identifying effective methods and approaches to improve ARDS prognosis. Early identification of high-risk ARDS patients and proactive intervention have been shown to mitigate the occurrence and progression of ARDS [4–6].

Gajic et al. constructed a predictor of clinical indicators of ARDS: the LIPS, which was validated in a multicenter study [7]. However, the positive predictive value (PPV) of LIPS was only 18%, thus limiting its clinical application. Other prediction models for ARDS (e.g., surgical lung injury prediction model (SLIP), emergency department lung injury prediction score (EDLIPS), sepsis ARDS prediction model, traumatic brain injury ARDS prediction model) have not been validated in clinical practice [8–11].

In this study, we observed the incidence of ARDS in high-risk patients. The LIPS score and other clinical parameters were used to develop and validate an ML-based ARDS prediction model for accurate patient identification.

Methods

Study population and diagnosis of ARDS

From December 2020 to August 2023, we prospectively enrolled 476 consecutive patients with identified risk factors for ARDS in the Department of Surgical Critical Care Medicine at the PUTH. Among these patients, 117 were ultimately diagnosed with ARDS. The diagnostic criteria for ARDS adhered to the 2012 Berlin definition [12]: new or worsening respiratory symptoms within 1 week attributed to known clinical triggers; bilateral diffuse exudates on chest imaging not entirely attributable to pleural effusion or lobar/lung collapse; respiratory failure induced by pulmonary edema not entirely attributable to cardiac failure or fluid overload; and an oxygenation index < 300.

Inclusion criteria comprised: (1) age ≥ 18 years; (2) presence of one or more risk factors for ARDS, including sepsis, shock, acute abdomen, pneumonia, pulmonary contusion, aspiration, high-risk trauma, or high-risk surgeries (such as spinal, abdominal, cardiac, and major vascular surgeries) [6]. Exclusion criteria encompassed: (1) patients with a pre-existing diagnosis of ARDS before initial assessment; (2) ICU admission duration shorter than 72 h; (3) patients deceased within 6 h of admission; (4) incomplete core information; and (5) abandonment of intensive care. Patients meeting any of these conditions were excluded from the study. Ultimately, 400 patients were included in our analysis (Fig. 1). The study protocol received approval from the Ethics Committee of the PUTH (Approval No. M2020278). Patient treatments adhered to the diagnosis and treatment protocols issued by the National Health Commission in real-time, and informed consent was obtained from patients or their families. Relevant clinical data were collected and analyzed with strict adherence to patient identity confidentiality protocols.

Data collection and processing

We utilized demographic characteristics, vital sign measurements, and laboratory data collected within the first 24 h after ICU admission to identify features and construct prediction models, as shown in Table 1. All data were extracted from the Electronic Medical Record (EMR) system. Given that nursing record system entries on the first day after ICU admission were documented from admission time until 7 am the subsequent day, we utilized the recorded data during this period for constructing the prediction model. The LIPS, Acute Physiology and Chronic Health Evaluation (APACHE) II score, Sequential Organ Failure Assessment(SOFA) score, and Neutrophil-to-Lymphocyte Ratio (NLR) were computed based on physiological parameters documented within the first 24 h of ICU admission.

To address multicollinearity, a test was performed to identify variables with a variance inflation factor (VIF) above 5, indicating multicollinearity (as shown in Table S1). A total of 35 features were retained, demographic factors including age, sex, body mass index (BMI); medical history including chronic obstructive pulmonary disease (COPD), hypertension, coronary artery disease (CAD), diabetes mellitus (DM), hepatic disease, malignancy, chronic kidney disease (CKD); use of vasoactive drugs, continuous renal replacement therapy (CRRT), or massive blood transfusion; admission type and diagnoses including shock, trauma, brain injury, sepsis, acute abdomen, pulmonary contusion, pneumonia, aspiration, orthopedic spinal, spinal cord injury (SCI); and clinical parameters and evaluations such as 24-h

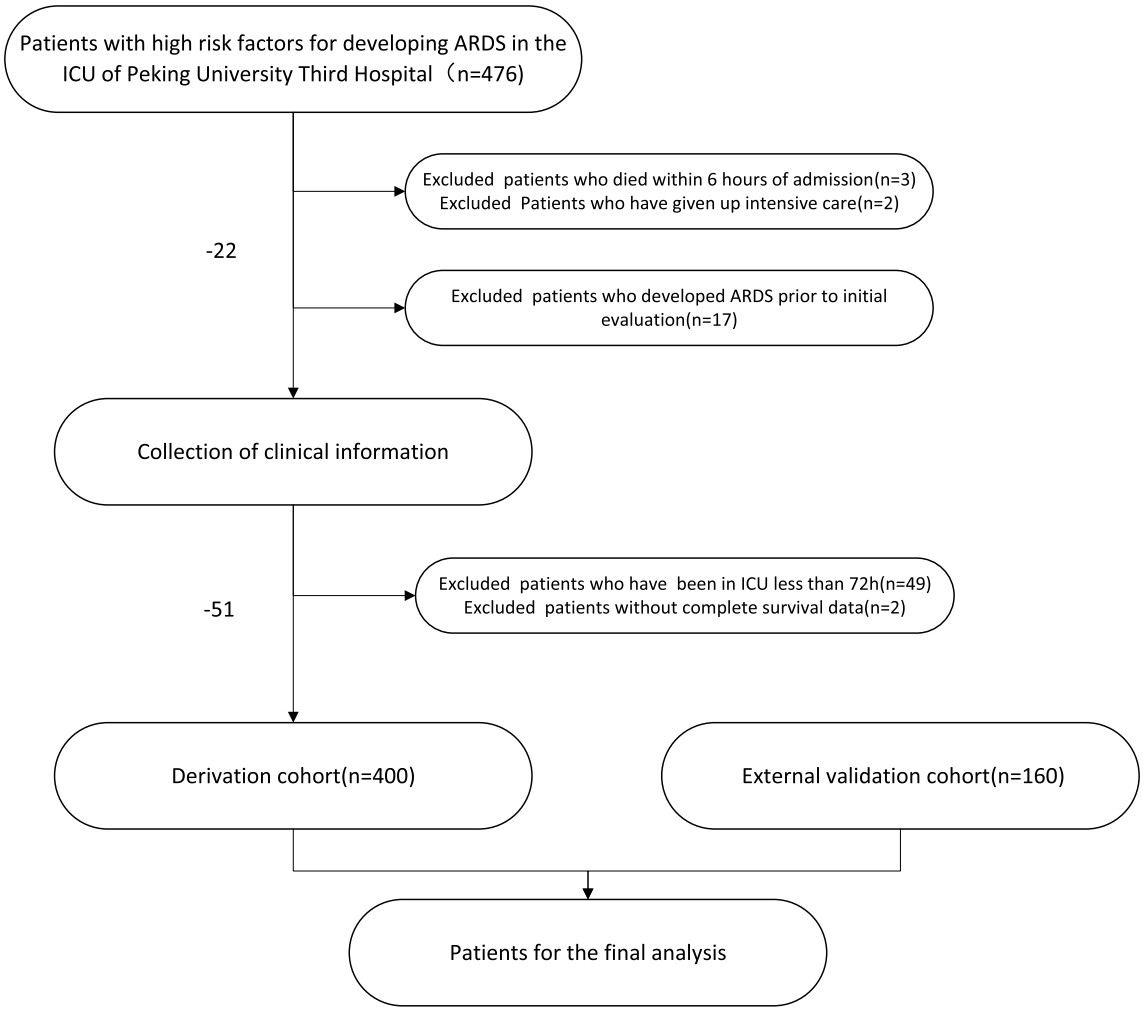

Fig. 1 Flowchart of screening

fluid balance upon admission (calculated as fluid intake minus fluid output), NLR, hemoglobin (HGB), red blood cell distribution width (RDW_CV), lactate, international normalized ratio (INR), urea, albumin (ALB), APACHE II scores, LIPS scores, and SOFA scores (Table 1).

External validation cohort

An external dataset comprising 160 patients from the ICU of FU YANG People's Hospital, collected between August 2022 and August 2023, was incorporated to validate the derivation cohort. The inclusion and exclusion criteria mirrored those applied to the training cohort.

Model development and comparison

First, R software (glmnet4.1.2) was used to conduct the least absolute shrinkage and selection operator (LASSO) regression analysis and adjust the variable screening and complexity. Subsequently, the outcomes of the LASSO regression analyses were utilized to conduct multifactor logistic regression analyses, ultimately identifying characteristic factors with P-values < 0.05. The Python software (version 0.22.1) was then employed to randomly partition the data from the PUTH into 70% for training and 30% for validation to mitigate potential overfitting issues. Additionally, an external dataset was employed for testing purposes (external validation).

The final screened features were used to develop predictive models. Five ML models namely K Nearest Neighbours (KNN), Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), and eXtreme Gradient Boosting (XGboost) were used for predicting the occurrence of ARDS in critically ill patients. To optimize the prediction model, grid search combined with manual fine-tuning was applied to obtain the final hyperparameters.

Several commonly used evaluation indexes, such as AUC, sensitivity, specificity, PPV, negative predictive value(NPV), accuracy, and F1 score, were used to

Table 1 Baseline demographics, clinical characteristics, and outcomes

	Non-ARDS (n = 283)	ARDS (n = 117)	*P* value
Age, (year), median (Q1,Q3)	62.0 (50.0, 73.0)	66.0 (55.0, 76.0)	0.070
High (m), median (Q1,Q3)	1.68 (1.60, 1.72)	1.70 (1.63, 1.73)	0.215
Weight (kg), median (Q1,Q3)	68.0 (60.0, 75.0)	69.0 (60.0, 76.0)	0.641
BMI (kg/m^2), median (Q1,Q3)	24.1 (21.5, 26.3)	23.9 (21.6, 26.7)	0.824
Gender (male), n (%)	159 (56.18)	86 (73.50)	0.001
Fluid intake	3084.0 (2703.5, 3991.0)	3413.00 (2788.0, 4346.0)	0.024
Fluid output	2800.00 (2032.50, 3601.00)	2608.00 (1915.00, 3460.00)	0.067
Fluid balance	2800.0 (2032.5, 3601.0)	2608.0 (1915.0, 3460.0)	< 0.01
Admission type, n (%)			
Planned surgical	87 (30.74)	21 (17.95)	0.024
E mergency surgical	141 (49.82)	65 (55.56)	
Medica	55 (19.43)	31 (26.50)	
Comorbidities, n (%)			
COPD	6 (2.12)	2 (1.71)	1.000
Hypertensive	129 (45.58)	60 (51.28)	0.299
CAD	30 (10.60)	19 (16.24)	0.118
DM	61 (21.55)	29 (24.79)	0.481
Hepatic disease	33 (11.66)	29 (24.79)	< 0.001
Tumor	39 (13.78)	24 (20.51)	0.093
CKD	12 (4.24)	7 (5.98)	0.456
Laboratory variables median (IQR)			
HGB	113.0 (94.0,128.0)	114.0 (92.0, 131.0)	0.952
NLR	9.8 [5.4,18.3]	9.6 [5.1,18.50]	0.991
RDW_CV	13.3 (12.7, 14.6)	13.9 (12.9, 14.8)	0.054
Lactate	2.0 (1.4, 3.4)	2.3 (1.5, 4.0)	0.084
INR	1.2 (1.1, 1.4)	1.3 (1.2 1.5)	0.009
UREA	6.7(4.8, 10.2)	8.7 (6.0, 12.6)	< 0.001
ALB	28.0 (23.8 32.1)	27.6 (23.3, 30.5)	0.319
Diagnose n (%)			
Shock	100 (35.34)	81 (69.23)	< 0.01
Trauma	36 (12.72)	19 (16.24)	0.353
Brain injury	67 (23.67)	20 (17.09)	0.147
Sepsis	97 (34.28)	65 (55.56)	< 0.01
Acute abdomen	102 (36.04)	64 (54.70)	< 0.001
Lung contusion	6 (2.12)	15 (12.82)	< 0.001
Pneumonia	21 (7.42)	24 (20.51)	< 0.001
Aspiration	0 (0.4)	7 (6.0)	< 0.001
Orthopedic spine	36 (12.72)	6 (5.13)	0.024
SCI	51 (18.02)	9 (7.69)	0.008
Advanced life support n (%)			
MBT	3 (1.06)	5 (4.27)	0.090
Vasoactive drugs	184 (65.02)	97 (82.91)	< 0.001
CRRT	38 (13.43)	33 (28.21)	< 0.001
Score median (Q1,Q3**)**			
APACHE II	17.0 (13.0, 19.0)	19.0 (16.0, 22.0)	< 0.001
LIPS	5.5 (3.5, 7.0)	8.5 (6.5, 10.0)	< 0.001
SOFA	6.0 (4.0, 8.0)	8.0 (6.0, 10.0)	< 0.001

ARDS: acute respiratory distress syndrome NLR: Neutrophil-to-Lymphocyte Ratio; APACHE: Acute Physiology and Chronic Health Evaluation; SOFA: Sequential Organ Failure Assessment; BMI: body mass index; COPD: chronic obstructive pulmonary disease; CAD: coronary artery disease; DM: diabetes mellitus; CKD: chronic kidney disease; MBT: massive blood transfusion; CRRT: continuous renal replacement therapy; SCI: spinal cord injury; HGB: hemoglobin; RDW_CV:Coefficient of variation of red cell distribution width; INR: international normalized ratio; ALB: albumin;

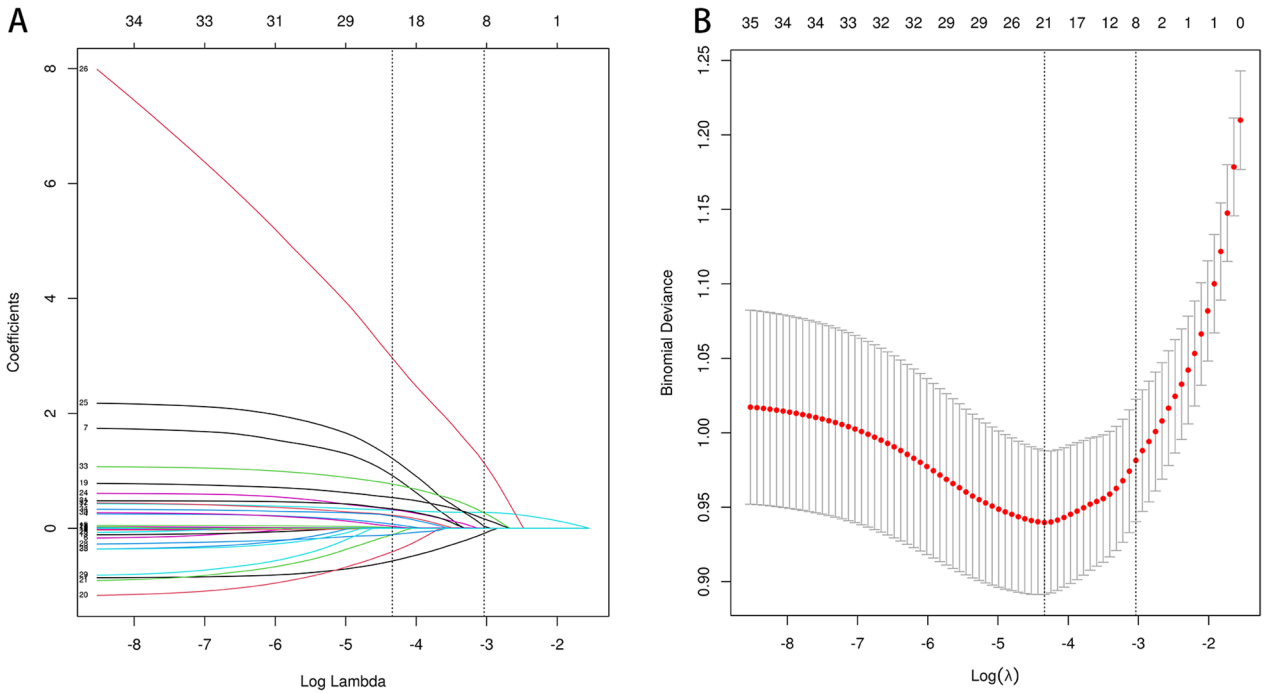

Fig. 2 Results of the LASSO regression analysis. **A** Plot of the LASSO coefficient profiles. **B** Tuning parameter (λ) selection cross-validation error curve

evaluate the reliability of these models. Furthermore, a DCA was carried out to evaluate the utility of the decision models by quantifying the net benefit across different threshold probabilities.

Model explanation

The SHapley Additive exPlanations (SHAP) method offered global and local explanations for the model explanation. The global explanation could give consistent and accurate attribution values for each feature within a model to show the associations between input features and ARDS. The local explanation could demonstrate a specific prediction for individual patients by inputting the specific data.

Webpage deployment tool based on shiny

To enhance the clinical utility of the model, a web-based interactive dynamic column chart application was developed using Shiny (version 0.13.2.26). This application facilitates the prediction of the probability of ARDS by inputting the values of the corresponding features from the final model.

Statistical analysis

The data were analyzed using Python version 3.6.5 (https://www.python.org) and SPSS version 27.0 (https://www.ibm.com/spss). Categorical data were presented as frequencies (percentages) and compared using chi-square tests. Continuous variables were presented as median (interquartile range) and compared using the Kruskal–Wallis test. The predictive power was evaluated using the AUC, with the optimal cut-off value determined by maximizing the Youden index (sensitivity + specificity − 1). DCA and Precision-Recall (P-R) curve analysis was conducted using R version 4.1.2 (https://www.r-project.org). A significance level of $P < 0.05$ was considered statistically significant.

Results

Patient characteristics

The study utilized 400 patients from the PUTH as the derivation cohort, with 117 patients diagnosed with ARDS, yielding an incidence rate of 29.3% during hospitalization. The derivation cohort was randomly divided into two subsets using the random number method: 70% of patients comprised the training cohort, while the remaining 30% constituted the internal validation cohort for constructing and screening the best predictive model. Additionally, an external cohort comprising 160 patients from FU YANG People's Hospital was employed for external validation. Details of the study design are displayed in Fig. 1

As shown in Table 1, among the 400 patients in the derivation cohort, a notably higher proportion of patients in the ARDS group compared to the non-ARDS group were male. Significant differences between the ARDS

Table 2 Multivariate logistic regression analysis

Predictor	R	SE	Z	*p*	Odds ratio
(Intercept)	− 6.052	1.397	− 4.333	0	
Admission type	0.248	0.239	1.039	0.299	1.281 [0.803, 2.053]
MBT	1.904	1.084	1.756	0.079	6.714 [0.925, 65.125]
Fluid balance	0	0	0.909	0.363	1 [1]
RDW_CV	0.012	0.062	0.192	0.848	1.012 [0.892, 1.139]
ALB	0.048	0.025	1.89	0.059	1.049 [0.999, 1.104]
APACHE II	0.008	0.031	0.268	0.788	1.008 [0.948, 1.073]
LIPS	0.378	0.073	5.154	0	1.459 [1.27, 1.694]
SOFA	− 0.016	0.052	− 0.302	0.763	0.984 [0.888, 1.09]
Gender (Female)	− 0.908	0.312	− 2.91	0.004	0.403 [0.215, 0735]
Shock (Yes)	0.722	0.35	2.062	0.039	2.059 [1.039, 4.123]
Trauma (Yes)	− 1.045	0.625	− 1.673	0.094	0.352 [0.095, 1.120]
Brain Injury (Yes)	− 0.586	0.438	− 1.337	0.181	0.557 [0.229, 1.29]
Pneumonia (Yes)	0.524	0.453	1.156	0.247	1.688 [0.687, 4.101]
Lung Contusion (Yes)	2.291	0.765	2.994	0.003	9.889 [2.325, 48.62]
Aspiration (Yes)	17.766	776.189	0.023	0.982	51,944,469
SCI (Yes)	− 0.195	0.583	− 0.335	0.738	0.823 [0.245, 2.472]
Hypertensive (Yes)	0.192	0.311	0.619	0.536	1.212 [0.66, 2.239]
CAD (Yes)	0.501	0.407	1.231	0.218	1.651 [0.738, 3.666]
DM (Yes)	0.404	0.354	1.141	0.254	1.498 [0.745, 2.999]
Hepatic Disease (Yes)	1.076	0.387	2.782	0.005	2.933 [1.383, 6.335]
Tumor (Yes)	0.421	0.395	1.066	0.286	1.524 [0.698, 3.308]

and non-ARDS groups were observed in terms of gender, fluid balance, liver function abnormality, and urea levels ($P < 0.05$). Furthermore, within the initial 24 h of admission to the ICU, patients in the ARDS group exhibited significantly higher LIPS scores, APACHE II scores, and SOFA scores ($P < 0.01$). Variables were compared between the derivation cohort and the external validation cohort (as shown in Table S2). Patients in the external validation cohort were older than those in the derivation cohort. Patients in the external validation cohort had higher LIPS scores, SOFA scores, APACHE II scores, and higher NLR and RDW_CV compared with the derivation cohort.

Factor selection for the predictive model

LASSO regression analysis was conducted on the remaining independent variables, with ARDS serving as the dependent variable (Fig. 2). LASSO can compress variable coefficients to prevent overfitting and solve severe collinearity problems [13]. The results showed (lambda = 0.013 for minimum mean square error) that 35 independent variables were reduced to 21, including gender, admission type, MBT, Fluid balance, RDW_CV, ALB, APACHE II, LIPS, SOFA, shock, Trauma, Brain Injury, Lung contusion, Pneumonia, Lung contusion, Aspiration, SCI, Hypertensive, CAD, DM, HepaticDisease, and Tumor. To further control for the effects of confounders, multiple logistic regression was used to analyze the above 21 independent variables were analyzed [14]. Finally, only Gender, LIPS, shock, lung contusion, and Hepatic disease were identified as characteristic factors ($P < 0.05$) (as Table 2).

Comprehensive analysis of classified multi-model

XGBoost, LR, RF, SVM, and KNN were trained and repeated 10 times. The models were evaluated using AUC values [15], which showed that XGBoost and RF were highest in the training set, and LR was highest in the set of internal validation sets (Fig. 3a–c). The AUC indicator focuses on the predictive accuracy of the model and does not tell whether the model is clinically usable or which one of the two is preferable [16]. Therefore, the DCA, calibration curves, and PR curves were analyzed. The DCA assessment revealed that the Logistic Regression model exhibited better clinical applicability (Fig. 3d). Calibration curves demonstrated higher accuracy in the predictions of the Logistic model (Fig. 3e). Moreover, in the validation set, the Logistic model demonstrated higher Average Precision (AP) values (Fig. 3f). Taken together, these findings suggest that the Logistic Regression model was the most desirable model for this study (Details are displayed in Tables S3 and S4).

The best model building and external evaluation

Logistic regression analysis and tenfold cross-validation were performed on the training set. The results show an average AUC of 0.824 (0.779–0.869) for the training set, 0.808 (0.662–0.951) for the internal validation set, and 0.799 (0.723–0.875) for the external validation set (Fig. 4a–c) (see Table S5 S6and S7 for more details). Given that the performance of the validation set under the AUC metric did not surpass that of the test set, or the exceedance ratio was less than 10%, the fit can be considered successful [16]. This demonstrates that the model exhibited strong performance in both internal and external validation. Additionally, the DCA assessment model

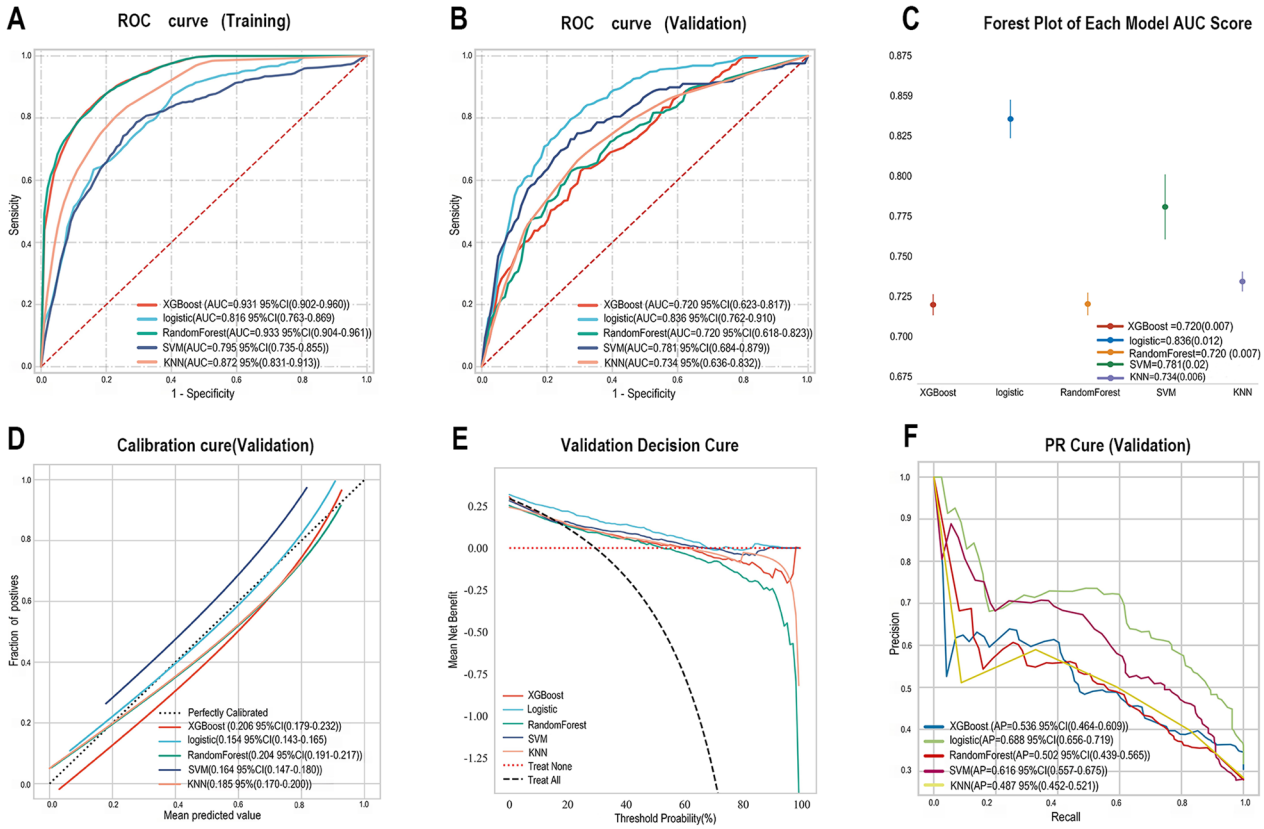

Fig. 3 ML model comprehensive analysis. **A** Training sets ROC and AUC **B** Validation sets ROC and AUC. Patients were sampled 10 times in a 7:3 ratio. **C** The validation set AUC forest plot **D** Calibration curves for the validation set, with the horizontal coordinates being the average predicted probability, the case coordinates being the actual probability of the event, the dashed diagonal being the reference line, and the rest of the solid lines representing the different models. **E** The validation set DCA where the black dotted line is the assumption that all patients are treated and the red represents the assumption that all patients are not treated. The remaining solid lines represent different models. **F** Validation set PR curve and AP The y-axis is precision and the x-axis is recall. If the PR curve of one model is completely covered by the PR curve of another model, it can be concluded that the latter is better than the former, and the higher the AP value, the better the model performance. The different colors in the picture represent the corresponding model, and the values are represented by the average and 95% CI

demonstrated good clinical applicability in the external test set (Fig. 4d).

Model explanation

To visually elucidate the selected variables, SHAP was utilized to illustrate how these variables predict the occurrence of ARDS in the model [17]. Figure 5a depicts the five most important features in our model. Each feature significance line showcases all patient attributions for the outcome, represented by differently colored dots: red dots indicate high-risk values, while blue dots denote low-risk values. Elevated LIPS score, male gender, shock, pulmonary contusion, and combined liver disease were found to increase the risk of ARDS in high-risk patients. Figure 5b displays the ranking of the five risk factors assessed by mean absolute SHAP value, with the x-axis SHAP value indicating the importance of the predictive model. Additionally, two typical examples are provided to illustrate the interpretability of the model. One example pertains to patients who developed ARDS with a higher SHAP prediction score (0.82) (Fig. 5c), while the other example involves patients who did not develop ARDS, exhibiting a lower SHAP prediction score (0.24) (Fig. 5d).

Convenient application for clinical utility

The final prediction model was integrated into a web application to enhance its utility in clinical settings, as depicted in Fig. 6. This application seamlessly predicts a patient's risk of developing ARDS upon entering the actual values of the five features required by the model. The web application is accessible online at https://prediction-model-for-ards.shinyapps.io/Prediction_ards/

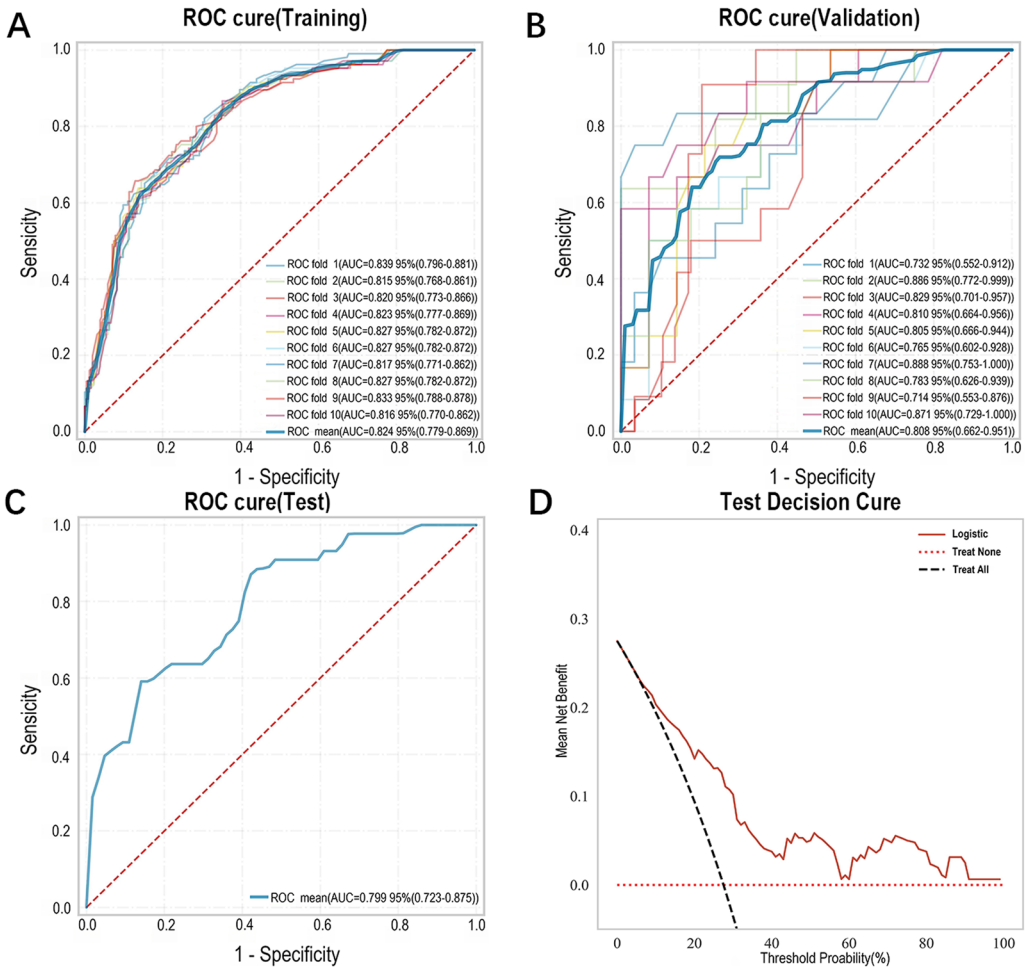

Fig. 4 Logistic regression model training, validation, and testing. **A** Training set ROC and AUC and **B** Validation set ROC and AUC. Training and cross-validation of patients in the derivation cohort. Different colored solid lines represent 10 different outcomes. **C** Test set ROC and AUC. test results from patients of the external test cohort. **D** Test set clinical decision curves, where the black dashed line represents the hypothesis that all patients have ARDS and the red represents the hypothesis that no patients have ARDS. The solid line represents the logistic model

Discussion

The findings of this study revealed that among ICU patients with high-risk factors for ARDS, the incidence of ARDS was 29.3%. In comparison, the LUNG SAFE study in 2016 reported that approximately 10% of ICU inpatients and 23.4% of mechanically ventilated patients developed ARDS [2]. Additionally, the CHARDS net study highlighted that the incidence rate of ARDS in 18 ICUs across mainland China was 3.57%. However, notable disparities were observed among different ICUs, with incidence rates varying significantly from one ICU to another, ranging from 1.0% to 16.7% [3]. The slightly higher incidence rate of ARDS observed in this study compared to the national study could potentially be attributed to the prospective observational nature of this study, which exclusively included patients with high-risk factors for ARDS meeting admission criteria, thus potentially skewing the incidence rate. Furthermore, differences in research methodologies and treatment management among various research institutions may also contribute to this discrepancy in incidence rates.

Despite advancements, the morbidity and mortality rates of ARDS patients remain substantial, reaching as high as 40.0%, with these rates escalating with the severity of ARDS [18]. However, clinical diagnosis of ARDS is frequently delayed or overlooked. In the Lung Safe study, only 60.2% of all ARDS patients were identified by clinicians, with recognition rates ranging from 51.3% for mild ARDS to 78.5% for severe ARDS [2]. Additionally, less than two-thirds of ARDS patients receive the recommended tidal volume of 8 mL/kg or less of their expected body weight [2].

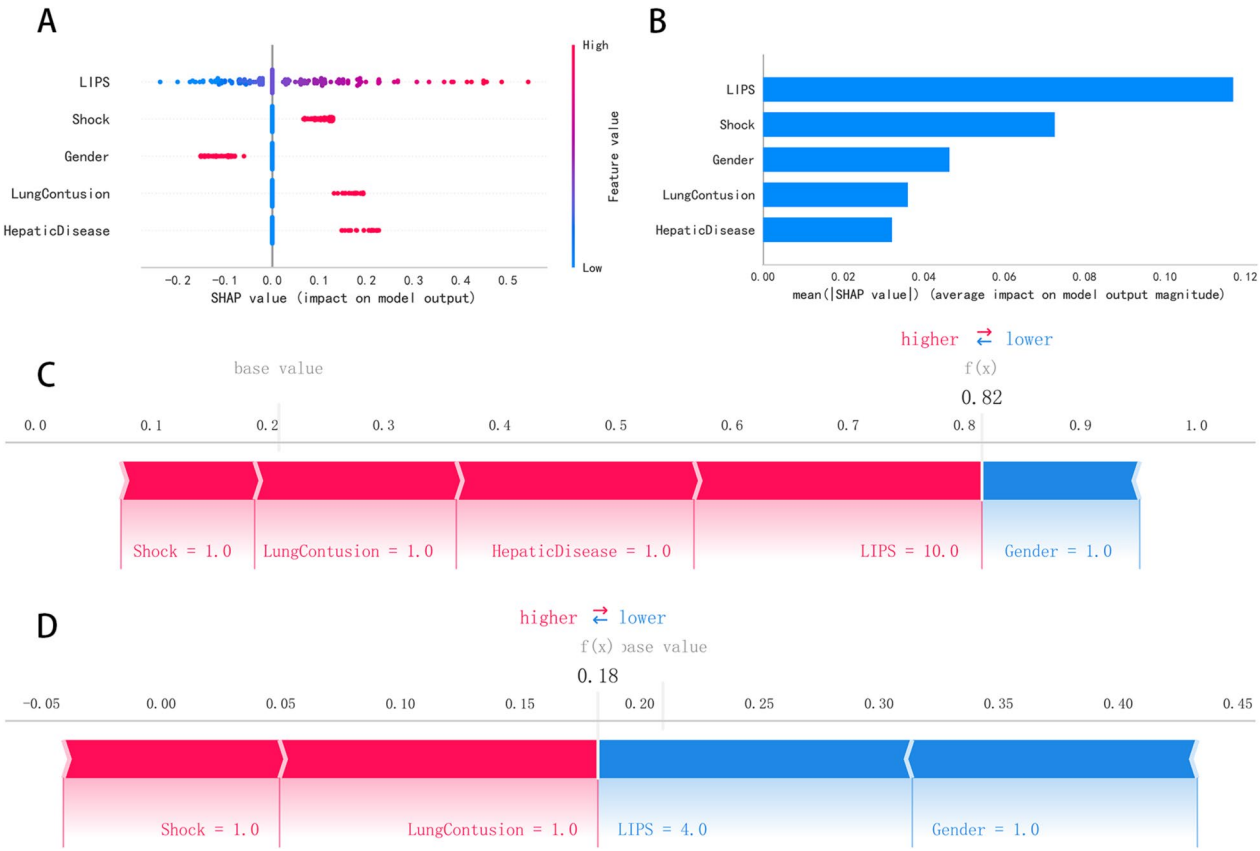

Fig. 5 SHAP interprets the mode. **A** Attributes of characteristics in SHAP. Each line represents a feature and the abscissa is the SHAP value. Red dots represent higher eigenvalues and blue dots represent lower eigenvalues. **B** Feature importance ranking as indicated by SHAP. The matrix diagram describes the importance of each covariate in the development of the final predictive model. **C** Individual efforts by patients with ARDS and **D** without ARDS. The SHAP value represents the predicted characteristics of an individual patient and the contribution of each characteristic to the predicted ARDS. The number in bold is the probability forecast value (f(x)), while the base value is the predicted value without providing input to the model. F(x) is the logarithmic ratio of each observation. Red features indicate an increased risk of ARDS and blue features indicate a reduced risk of ARDS. The length of the arrows helps visualize the extent to which the prediction is affected. The longer the arrow, the greater the effect

Delayed recognition of ARDS can fail to implement strategies known to improve ARDS survival, such as protective mechanical ventilation [19], fluid restriction [20], and prone ventilation [21, 22], consequently leading to increased mortality [23]. Despite ongoing research, there remains a lack of a widely accepted predictive model to identify individuals at risk of ARDS. One notable strength of this study is its prospective design, wherein patients were screened daily until the 7th day of ICU admission, effectively mitigating missed or misdiagnosed cases of ARDS. Moreover, by employing well-established ARDS risk prediction models and disease severity scoring systems from previous studies, this study has developed a prediction model for ARDS incidence that aligns with clinical practice and exhibits improved general applicability.

In this study, we employed LASSO regression and multifactorial logistic regression analyses to identify five key variables (Gender, LIPS, shock, lung contusion, and Hepatic disease) out of 35 factors associated with the risk of ARDS in patients. The LIPS score model, proposed by Gajic and Trillo-Alvarez in 2011, has been widely used in recent years to assess the risk stratification of patients with acute lung injury (ALI). It systematically grades the extent and degree of lung injury based on factors such as susceptibility, high-risk surgeries and traumas, and other risk-modifying factors. The LIPS score has demonstrated good predictive value for ALI, with AUC ranging from 0.80 to 0.84 [7, 24]. Studies conducted in China by Xiejianfeng et al. reported an AUC of 0.770 for the LIPS score [25], while Soto et al. demonstrated a close association between the LIPS score and the occurrence of ARDS, with an AUC of 0.740 [26]. The sensitivity of the LIPS score ≥ 4 in predicting ALI was 90.3%, the specificity was 30.9%, the positive predictive value was 17.3%, and the negative

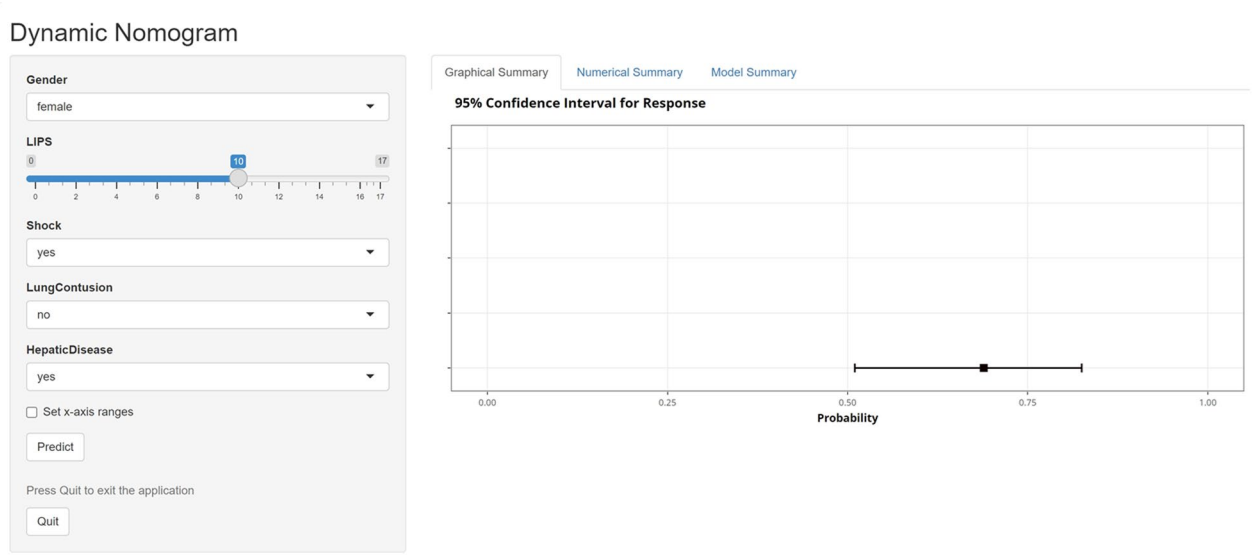

Fig. 6 Online dynamic nomogram. An illustrative example is described for predicting the probability of ARDS after ICU admission in a woman presenting in shock with comorbid liver disease with a LIPS score of 10

predictive value was 95.2% [26]. The specificity of the LIPS score in predicting the occurrence of ARDS is low, and it may not be accurate for clinical risk prediction alone. Our model combines the LIPS score with other clinical data, and the final model had a sensitivity of 63.6%, a specificity of 83.6%, a positive predictive value of 41.9%, and a negative predictive value of 92.5% in the test set. (Table S8).

Furthermore, male patients are more predisposed to ARDS compared to females. In a secondary analysis of the LUNG SAFE study, it was found that ARDS occurred more commonly in men than in women, with 905 (38%) women compared to 1472 (62%) men out of 2377 ARDS patients [27]. A similar male predisposition to ARDS has been reported in patients with COVID-19-associated ARDS [28]. Patients with liver disease are also at a higher risk of developing ARDS [29]. In patients with COVID-19, advanced liver disease has been associated with worse respiratory outcomes, including increased mortality and the need for mechanical ventilation [30]. This could be attributed to impaired function of the hepatic reticuloendothelial system and hepatocytes, which compromises systemic and pulmonary defenses, thereby impacting the onset and remission of ARDS [31, 32]. In our study, combined liver disease was identified as an independent risk factor for the development of ARDS. Additionally, shock and pulmonary contusion were identified as risk factors for ARDS in both the Lung Safe study and our study [2]. However, other risk factors such as sepsis, pneumonia, and trauma did not show predictive value for the development of ARDS in our multifactorial logistic regression analysis, possibly due to covariance with factors included in the LIPS scoring system.

In this study, we used several ML models and found that logistic regression models outperformed other models after analyzing AUC, DCA, calibration curves, and PR curves. However, it has been a challenge to interpret machine learning predictive models more comprehensively and to present the predictions visually to clinicians. Therefore, the SHAP approach was used to interpret the logistic model, which ensures model performance and clinical interpretability. This will help physicians better understand the decision-making process of the model and facilitate the use of the prediction results. In the external validation cohort, when the prediction specificity was fixed at 83.6%, the negative predictive value was 92.5% and the positive predictive value was only 41.9%. Therefore, the model may not be able to fully provide decision support to clinicians. In clinical practice, it is necessary to evaluate the benefits of early identification of people at risk of ARDS and their additional costs.

Indeed, our study has several limitations that warrant consideration. Firstly, it is a prospective single-center study with a relatively small sample size. Consequently, the generalizability of our findings may be limited, and validation in larger, multicenter studies is necessary to confirm the robustness of our results. Secondly, while we validated our model using an external cohort, it is essential to acknowledge that this cohort consisted of retrospective data. This may introduce selection bias and affect the generalizability of our model's performance to other settings. Thirdly, our prediction model was based

on data available within 24 h of ICU admission and failed to use data on dynamic changes in indicators, which may have somewhat overlooked subsequent events that altered the occurrence of ARDS and caused confounding factors.

Conclusions

In conclusion, our ML model, incorporating sex, LIPS, shock, pulmonary contusion, and liver disease as risk factors, effectively predicts ARDS in high-risk patients. Compared to the LIPS score, the model demonstrates a significantly higher positive predictive value (41.9% vs. 18%). In high-risk populations, the model enhances the identification of truly at-risk individuals, optimizes medical resource allocation, and provides additional intervention time for patients. Furthermore, the SHAP method was used to explain personalized ARDS risk, fostering greater trust among clinicians in the model's predictions. However, these conclusions require validation through further randomized controlled trials.

Abbreviations

ARDS	Acute respiratory distress syndrome
ICU	Intensive care units
LASSO	Least absolute shrinkage and selection operator
DCA	Decision curve analysis
ROC	Receiver operating characteristic
AUC	Area under the receiver operating characteristic curve
ML	Machine learning
LIPS	Lung injury prediction score
SHAP	S Hapley Additive ex Planations
PPV	Positive predictive value
SLIP	Surgical lung injury prediction model
EDLIPS	Emergency department lung injury prediction score
PUTH	Peking University Third Hospital
EMR	Electronic medical record
NLR	Neutrophil-to-lymphocyte ratio
APACHE	Acute physiology and chronic health evaluation
SOFA	Sequential organ failure assessment
VIF	Variance inflation factor
BMI	Body mass index
COPD	Chronic obstructive pulmonary disease
CAD	Coronary artery disease
DM	Diabetes mellitus
CKD	Chronic kidney disease
MBT	Massive blood transfusion
CRRT	Continuous renal replacement therapy
SCI	Spinal cord injury
HGB	Hemoglobin
RDW_CV	Coefficient of variation of red cell distribution width
INR	International normalized ratio
ALB	Albumin
KNN	K nearest neighbours
LR	Logistic regression
RF	Random forest
SVM	Support vector machine
XGboost	EXtreme gradient boosting
NPV	Negative predictive value
P-R	Precision-recall
AP	Average precision
ALI	Acute lung injury

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-024-06054-1.

Additional file 1.

Acknowledgements

We are grateful to all the participants in the study for their contributions.

Author contributions

XZ, Q-GG, and D-FL had full access to all of the data in the study and took responsibility for the integrity of the data, the accuracy of the data analysis, and the administrative, technical, or material support. XZ, Q-GG, D-FL, S-HW, and HZ contributed to the study concept and design; HL, CL, Z-YS, Z-KC, Y-YY, Z-JZand S-HW contributed to the acquisition of data; S-HW HZ, CL, and HL contributed to the drafting of the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by Capital Health Research and Development of Special Fund (Grant/Award 20202-4094), Peking University Third Hospital Cohort Study Project (BYSYDL2021010), Peking University Third Hospital Cohort Study Project (BYSYZD2023012), National Natural Science Foundation of China (Grant/Award 82172166), Beijing Municipal Natural Science Foundation (Grant/Award 7212130).

Availability of data and materials

The data supporting the conclusions of this study are available from the corresponding authors upon reasonable request.

Declarations

Ethics approval and consent to participate

Ethics approval was acquired from the institutional review boards of all participating institutions, and written informed consent was obtained from all participants at enrollment.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Critical Care Medicine, Peking University Third Hospital, Beijing 100191, China. [2]Clinical Epidemiology Research Center, Peking University Third Hospital, Beijing 100191, China. [3]Department of Critical Care Medicine, Fuyang People's Hospital, Fuyang 236000, China.

Received: 29 March 2024 Accepted: 25 December 2024
Published online: 13 January 2025

References

1. Matthay MA, Arabi Y, Arroliga AC, Bernard G, Bersten AD, Brochard LJ, Calfee CS, Combes A, Daniel BM, Ferguson ND, et al. A new global definition of acute respiratory distress syndrome. Am J Respir Crit Care Med. 2024;209:37–47.
2. Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, Gattinoni L, van Haren F, Larsson A, McAuley DF, et al. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. JAMA. 2016;315:788–800.
3. Huang X, Zhang R, Fan G, Wu D, Lu H, Wang D, Deng W, Sun T, Xing L, Liu S, et al. Incidence and outcomes of acute respiratory distress syndrome in intensive care units of mainland China: a multicentre prospective longitudinal study. Crit Care. 2020;24:515.

4. Festic E, Carr GE, Cartin-Ceba R, Hinds RF, Banner-Goodspeed V, Bansal V, Asuni AT, Talmor D, Rajagopalan G, Frank RD, et al. Randomized clinical trial of a combination of an inhaled corticosteroid and beta agonist in patients at risk of developing the acute respiratory distress syndrome. Crit Care Med. 2017;45:798–805.
5. Curley GF, Laffey JG, Zhang H, Slutsky AS. Biotrauma and ventilator-induced lung injury: clinical implications. Chest. 2016;150:1109–17.
6. Gong MN, Thompson BT. Acute respiratory distress syndrome: shifting the emphasis from treatment to prevention. Curr Opin Crit Care. 2016;22:21–37.
7. Gajic O, Dabbagh O, Park PK, Adesanya A, Chang SY, Hou P, Anderson H 3rd, Hoth JJ, Mikkelsen ME, Gentile NT, et al. Early identification of patients at risk of acute lung injury: evaluation of lung injury prediction score in a multicenter cohort study. Am J Respir Crit Care Med. 2011;183:462–70.
8. Kor DJ, Warner DO, Alsara A, Fernández-Pérez ER, Malinchoc M, Kashyap R, Li GX, Gajic O. Derivation and diagnostic accuracy of the surgical lung injury prediction model. Anesthesiology. 2011;115:117–28.
9. Elie-Turenne M-C, Hou PC, Mitani A, Barry JM, Kao EY, Cohen JE, Frendl G, Gajic O, Gentile NT, Illness USC, Injury Trials Group: Lung Injury Prevention Study I. Lung injury prediction score for the emergency department: first step towards prevention in patients at risk. Int J Emerg Med. 2012;5:33.
10. Jiang ZZ, Liu LP, Du L, Lv SS, Liang F, Luo YW, Wang CJ, Shen Q. Machine learning for the early prediction of acute respiratory distress syndrome (ARDS) in patients with sepsis in the ICU based on clinical data. Heliyon. 2024;10:e28143.
11. Wang RR, Cai LR, Zhang J, He M, Xu JG. Prediction of acute respiratory distress syndrome in traumatic brain injury patients based on machine learning algorithms. Medicina-Lithuania. 2023; 59.
12. Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, Camporota L, Slutsky AS. Acute respiratory distress syndrome: the Berlin Definition. JAMA. 2012;307:2526–33.
13. Muthukrishnan R, Rohini R. LASSO: A Feature Selection Technique In Predictive Modeling For Machine Learning. In IEEE International Conference on Advances in Computer Applications (ICACA); Oct 24; Bharathiar Univ, Coimbatore, INDIA. 2016: 18–20.
14. Shi AX, Zivich PN, Chu HT. A comprehensive review and tutorial on confounding adjustment methods for estimating treatment effects using observational data. Appl Sci-Basel. 2024;14:3662.
15. Janssens A, Martens FK. Reflection on modern methods: revisiting the area under the ROC Curve. Int J Epidemiol. 2020;49:1397–403.
16. Muschelli J. ROC and AUC with a binary predictor: a potentially misleading metric. J Classif. 2020;37:696–708.
17. Zhou D, Jiang Y, Zhong X, Cox NJ, Liu C, Gamazon ER. A unified framework for joint-tissue transcriptome-wide association and Mendelian randomization analysis. Nat Genet. 2020;52:1239–46.
18. Banavasi H, Nguyen P, Osman H, Soubani AO. Management of ARDS—what works and what does not. Am J Med Sci. 2021;362:13–23.
19. Karageorgos V, Proklou A, Vaporidi K. Lung and diaphragm protective ventilation: a synthesis of recent data. Expert Rev Respir Med. 2022;16:375–90.
20. Lee J, Corl K, Levy MM. Fluid therapy and acute respiratory distress syndrome. Crit Care Clin. 2021;37:867–75.
21. Rampon GL, Simpson SQ, Agrawal R. Prone positioning for acute hypoxemic respiratory failure and ARDS: a review. Chest. 2023;163:332–40.
22. Guérin C, Reignier J, Richard JC, Beuret P, Gacouin A, Boulain T, Mercier E, Badet M, Mercat A, Baudin O, et al. Prone positioning in severe acute respiratory distress syndrome. N Engl J Med. 2013;368:2159–68.
23. Bellani G, Pham T, Laffey JG. Missed or delayed diagnosis of ARDS: a common and serious problem. Intensive Care Med. 2020;46:1180–3.
24. Trillo-Alvarez C, Cartin-Ceba R, Kor DJ, Kojicic M, Kashyap R, Thakur S, Thakur L, Herasevich V, Malinchoc M, Gajic O. Acute lung injury prediction score: derivation and validation in a population-based sample. Eur Respir J. 2011;37:604–9.
25. Xie J, Liu L, Yang Y, Yu W, Li M, Yu K, Zheng R, Yan J, Wang X, Cai G, et al. A modified acute respiratory distress syndrome prediction score: a multicenter cohort study in China. J Thorac Dis. 2018;10:5764–73.
26. Soto GJ, Kor DJ, Park PK, Hou PC, Kaufman DA, Kim M, Yadav H, Teman N, Hsu MC, Shvilkina T, et al. Lung injury prediction score in hospitalized patients at risk of acute respiratory distress syndrome. Crit Care Med. 2016;44:2182–91.
27. McNicholas BA, Madotto F, Pham T, Rezoagli E, Masterson CH, Horie S, Bellani G, Brochard L, Laffey JG. Demographics, management and outcome of females and males with acute respiratory distress syndrome in the LUNG SAFE prospective cohort study. Eur Respir J. 2019;54:1900609.
28. Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. Considering how biological sex impacts immune responses and COVID-19 outcomes. Nat Rev Immunol. 2020;20:442–7.
29. Damm TW, Kramer DJ. The liver in critical illness. Crit Care Clin. 2016;32:425–38.
30. Zmudka K, Jaroszewicz J, Zarebska-Michaluk D, Rogalska M, Czupryna P, Rorat M, Kozielewicz D, Maciukajc J, Kiciak S, Krepa M, et al. Association between liver damage and disease progression markers with mortality risk and mechanical ventilation in hospitalized COVID-19 patients: a nationwide retrospective SARSTer study. Viruses-Basel. 2024;16:1530.
31. Herrero R, Sánchez G, Asensio I, López E, Ferruelo A, Vaquero J, Moreno L, de Lorenzo A, Bañares R, Lorente JA. Liver-lung interactions in acute respiratory distress syndrome. Intensive Care Med Exp. 2020; 8.
32. Perez Ruiz de Garibay A, Kortgen A, Leonhardt J, Zipprich A, Bauer M. Critical care hepatology: definitions, incidence, prognosis and role of liver failure in critically ill patients. Crit Care. 2022;26:289.

Publisher's Note