Cite as: Xiong YX, Mao YQ, Gu QK, Zhao MY, Meng XL. Identification and validation of immune-related biomarkers for sepsis-induced acute respiratory distress syndrome [J]. Chin J Clin Res, 2026, 39(5): 669-676.

DOI: 10.13429/j.cnki.cjcr.2026.05.004

Identification and validation of immune-related biomarkers for sepsis-induced acute respiratory distress syndrome

XIONG Yaxin* , MAO Yuqi, GU Quankuan, ZHAO Mingyan, MENG Xianglin

* *Department of Critical Care Medicine, First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang 150000, China*

Corresponding author: MENG Xianglin, E-mail: mengxianglin@hrbmu.edu.cn

Abstract: **Objective** To provide a scientific basis for understanding the disease mechanism and achieving precise diagnosis and treatment of sepsis-induced acute respiratory distress syndrome (ARDS) related to immune imbalance by identifying immune regulation-related differentially expressed genes (IRDEGs). **Methods** Based on sepsis alone and sepsis -induced ARDS samples gained from the Gene Expression Omnibus (GEO) database, IRDEGs were screened and subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. Weighted gene co-expression network analysis (WGCNA) was used to obtain co-expression module genes. Combined with logistic regression, random forest, and LASSO regression, an ARDS diagnostic model was constructed. Meanwhile, CIBERSORT was employed to analyze immune infiltration. Peripheral blood samples from patients withsepsis-induced ARDS and sepsis alone (n=10, each), treated at the First Affiliated Hospital of Harbin Medical University in 2024, were collected to validate the protein expression of model genes via enzyme-linked immunosorbent assay (ELISA). **Results** A total of 34 IRDEGs and 4 model genes [tumor necrosis factor superfamily member 10 (*TNFSF*10), β2-microglobulin (B2M), semaphorin 7A (*SEMA7A*), and cluster of differentiation 200 (CD200)] were identified. An effective diagnostic model for ARDS was subsequently constructed. Immune infiltration analysis revealed extensive enrichment of immune cells and significant correlations in ARDS, such as a strong positive correlation between follicular helper T cells and resting dendritic cells (r=0.765, P<0.05). Model genes were associated with immune cell infiltration; for example, *B2M* was negatively correlated with naive B cell (r=-0.383, P<0.05). ELISA confirmed that serum levels of CD200 and SEMA7A were significantly higher in patients with sepsis -induced ARDS compared to those with sepsis alone (P<0.01). **Conclusion** The ARDS diagnostic model constructed using TNFSF10, B2M, SEMA7A, and CD200 exhibit good performance. There are significant correlations between the model genes and immune cells, as well as among various immune cells themselves. ELISA confirms that serum levels of CD200 and SEMA7A are significantly elevated in patients with septic ARDS, suggesting that they are potential diagnostic biomarkers.

Keywords: Acute respiratory distress syndrome; Sepsis; Immune regulation; Diagnostic model

Fund program: Key Research and Development Plan of Heilongjiang Province (GA21C011, JD22C005)

Acute respiratory distress syndrome (ARDS) is acute diffuse lung injury caused by multiple etiologies, which can lead to severe hypoxemia[1]. Sepsis is a common predisposing factor for ARDS, and sepsis-induced ARDS is associated with higher morbidity and mortality[2].

During the pathological process of sepsis-induced ARDS, immune cells are persistently activated and drive inflammatory cascades, releasing inflammatory mediators, which leads to the infiltration of peripheral immune cells into the lung parenchyma, damages the alveolar-capillary barrier, and causes diffuse pulmonary edema[3]. Immunoregulation involves changes in multiple immune cell subsets and molecules, and is an important research direction for ARDS. Studies have shown that the balance of M1/M2 phenotypic polarization of macrophages regulates the inflammatory process during acute lung injury (ALI)[4], and neutrophils, innate lymphoid cells and dendritic cells also play important roles in the occurrence and development of sepsis-induced ARDS[5-6]. Clinical studies have confirmed that biomarkers such as interleukin-6, interleukin-8 and pulmonary surfactant protein D have certain predictive value for the prognosis of ARDS patients[7-8]. At present, the immune molecular network of this disease has not been fully elucidated, and the identification of novel biomarkers is of great significance for optimizing diagnosis and treatment strategies. In recent years, gene expression characteristics have become the focus of disease research. Sun *et al.*[9] identified 7 autophagy-related genes, which can effectively distinguish sepsis-induced ARDS from simple sepsis. In another study, 4 senescence-related genes were found to have significant diagnostic value for sepsis-induced ARDS[10]. In addition, the expression of neutrophil-related genes is also significantly elevated in ARDS patients[11]. These gene signatures associated with specific biological processes are expected to serve as novel diagnostic markers and mechanistic research tools for sepsis-induced ARDS.

This study identified immune regulation-related differentially expressed genes (IRDEGs) in ARDS through bioinformatics analysis, and constructed an ARDS diagnostic model based on related genes [tumor necrosis

factor ligand superfamily member 10 (*TNFSF*10), beta-2-microglobulin (*B2M*), semaphorin 7A (*SEMA7A*) and cluster of differentiation 200 (CD200)]. Meanwhile, this study analyzed the difference in immune infiltration between sepsis-induced ARDS and simple sepsis using the cell-type identification by estimating relative subsets of RNA transcripts (CIBERSORT) algorithm. The expression levels of the above genes in the two groups of patients were further verified by enzyme-linked immunosorbent assay (ELISA), in order to provide a scientific basis for understanding the disease mechanism and realizing the precise diagnosis and treatment of ARDS.

1 Materials and Methods

1.1 Data sources

The datasets GSE32707 (including 31 sepsis-induced ARDS samples and 58 simple sepsis samples) and GSE66890 (including 29 sepsis-induced ARDS samples and 28 simple sepsis samples) were obtained from the Gene Expression Omnibus (GEO) database[12]. After batch effect removal using the R package "sva", the two datasets were merged to obtain a comprehensive dataset containing 60 ARDS samples and 86 control samples. Meanwhile, 237 immune regulation-related genes (IRGs) with a correlation score > 1 were screened from the GeneCards database.

1.2 IRDEGs

Differential expression analysis was performed between the two groups of samples to screen differentially expressed genes (DEGs), with the threshold set as $|\log_2$ fold change (FC)$| > 0.5$ and $P < 0.05$. The intersection of DEGs and IRGs was taken to obtain IRDEGs. Subsequently, the Spearman method was used to analyze the correlation between the expression levels of these genes.

1.3 Functional enrichment analysis

The clusterProfiler package was used to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis on IRDEGs. The screening criteria were adjusted P value < 0.05 and false discovery rate (FDR, i.e., q value) < 0.25. The Benjamini-Hochberg (BH) method was used for P value correction.

1.4 Construction of immune regulation score (I.Score) and weighted gene co-expression network analysis (WGCNA)

The single-sample gene-set enrichment analysis (ssGSEA) algorithm was used to calculate the I.Score of all samples based on the IRDEGs and the expression matrix of the comprehensive dataset. WGCNA was performed on the top 70% genes with the highest expression variance in the ARDS group to obtain co-expression modules. Modules significantly correlated with I.Score ($|r| > 0.40$) were screened, and the genes within these modules were intersected with IRDEGs to obtain module genes.

1.5 Construction and validation of ARDS diagnostic model

Logistic regression, support vector machine, and least absolute shrinkage and selection operator (LASSO) regression were comprehensively applied to screen module genes, and key model genes were obtained to construct the ARDS diagnostic model. The gene weights were evaluated through a nomogram, and the model calibration and clinical net benefit were evaluated by combining calibration curves and decision curve analysis, respectively. The RiskScore was calculated based on the LASSO risk coefficient, and the diagnostic efficacy of RiskScore was evaluated by the receiver operating characteristic (ROC) curve and the area under the curve (AUC).

1.6 Gene-set enrichment analysis (GSEA) of high- and low-risk groups

ARDS samples were divided into high-risk group and low-risk group based on the median RiskScore, and differential gene analysis was performed. The canonical pathway gene sets were obtained from the MSigDB database, and GSEA enrichment analysis was performed on the two groups of samples. The screening criteria were adjusted P value < 0.05 and FDR (q value) < 0.25, and the BH method was used for P value correction.

1.7 Immune infiltration analysis

The CIBERSORT algorithm combined with the leukocyte gene signature matrix 22 (LM22) was used to obtain the immune cell infiltration profile of the comprehensive dataset. Subsequently, Spearman correlation analysis was performed to evaluate the correlation between different immune cell types and the correlation between immune cell types and the expression of model genes.

1.8 Immune infiltration analysis of high- and low-risk groups

The ssGSEA was used to calculate the infiltration score based on immune cell subtypes, and an immune infiltration matrix was constructed. The differential expression of immune cells between the high- and low-risk groups was analyzed, and the correlation between immune cells and the correlation between immune cells and model genes were analyzed based on Spearman correlation analysis.

1.9 Clinical trial validation

Peripheral blood samples were collected from 10 patients with sepsis-induced ARDS (sepsis-ARDS group) and 10 patients with simple sepsis (sepsis group) who were admitted to the First Affiliated Hospital of Harbin Medical University in 2024. Blood samples were collected within 24 hours of sepsis diagnosis. After anticoagulation and centrifugation, serum was collected, and the levels of TNFSF10, B2M, SEMA7A and CD200 were quantitatively detected using ELISA kits. This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the First Affiliated Hospital of Harbin Medical University (No. IRB-AF/SC-04/0.2-0), and all subjects signed informed consent forms.

1.10 Statistical analysis

R 4.2.3 was used for statistical analysis and plotting. Independent sample *t* test was used for comparison between two groups, and one-way analysis of variance was used for comparison between multiple groups.

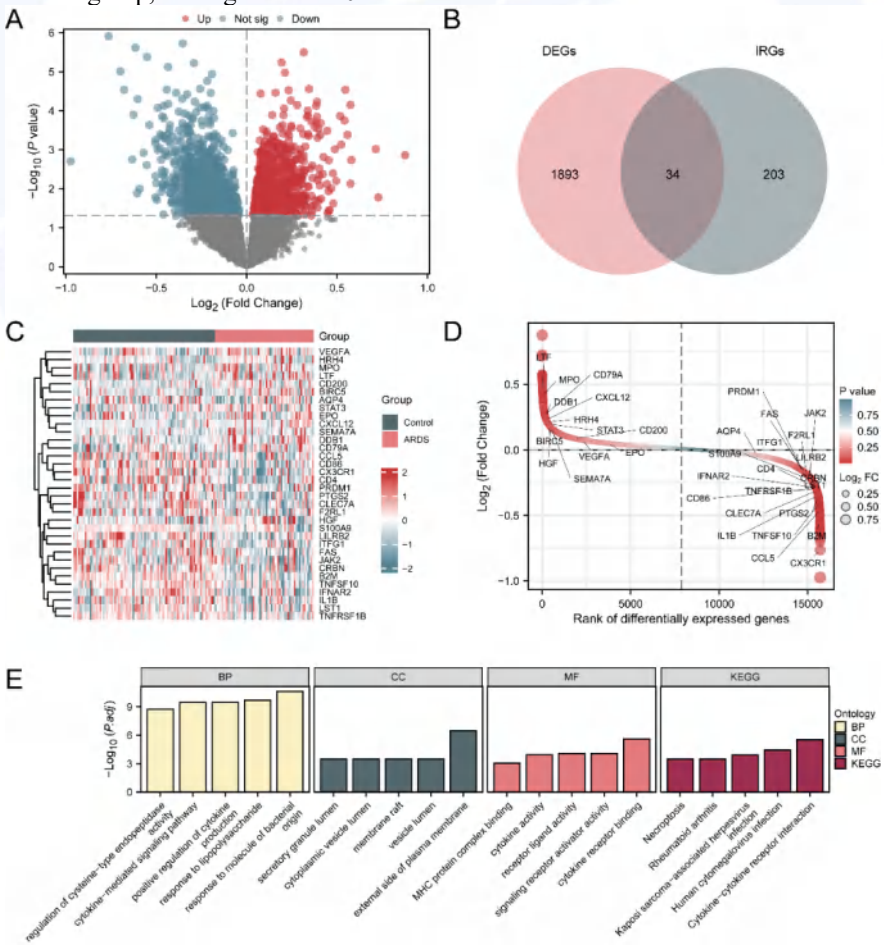
2 Results

2.1 ARDS IRDEGs

A total of 1 927 DEGs were identified between the ARDS group and the control group, among which 1 021 were upregulated and 906 were downregulated (Figure 1A). After intersecting with IRGs, 34 IRDEGs were obtained (Figure 1B). The expression differences of these genes between the two groups were visualized by heatmap and volcano plot (Figure 1C, Figure 1D).

2.2 Functional enrichment analysis

GO and KEGG enrichment analysis was performed on IRDEGs. The results showed that the biological processes were mainly enriched in response to molecules of bacterial origin, response to lipopolysaccharide, positive regulation of cytokine production, cytokine-mediated signaling pathway, and regulation of cysteine-type endopeptidase activity. For cellular components, the enrichment terms included external side of plasma membrane, lumen of secretory granule, cytoplasmic vesicle lumen, membrane raft, and vesicle lumen. For molecular functions, the enrichment terms included cytokine receptor binding, receptor ligand activity, signaling receptor activator activity, cytokine activity, and major histocompatibility complex (MHC) protein complex binding. Meanwhile, these genes were also enriched in the following KEGG biological pathways: cytokine-cytokine receptor interaction, human cytomegalovirus infection, Kaposi sarcoma-associated herpesvirus infection, necroptosis, and rheumatoid arthritis (Figure 1E).

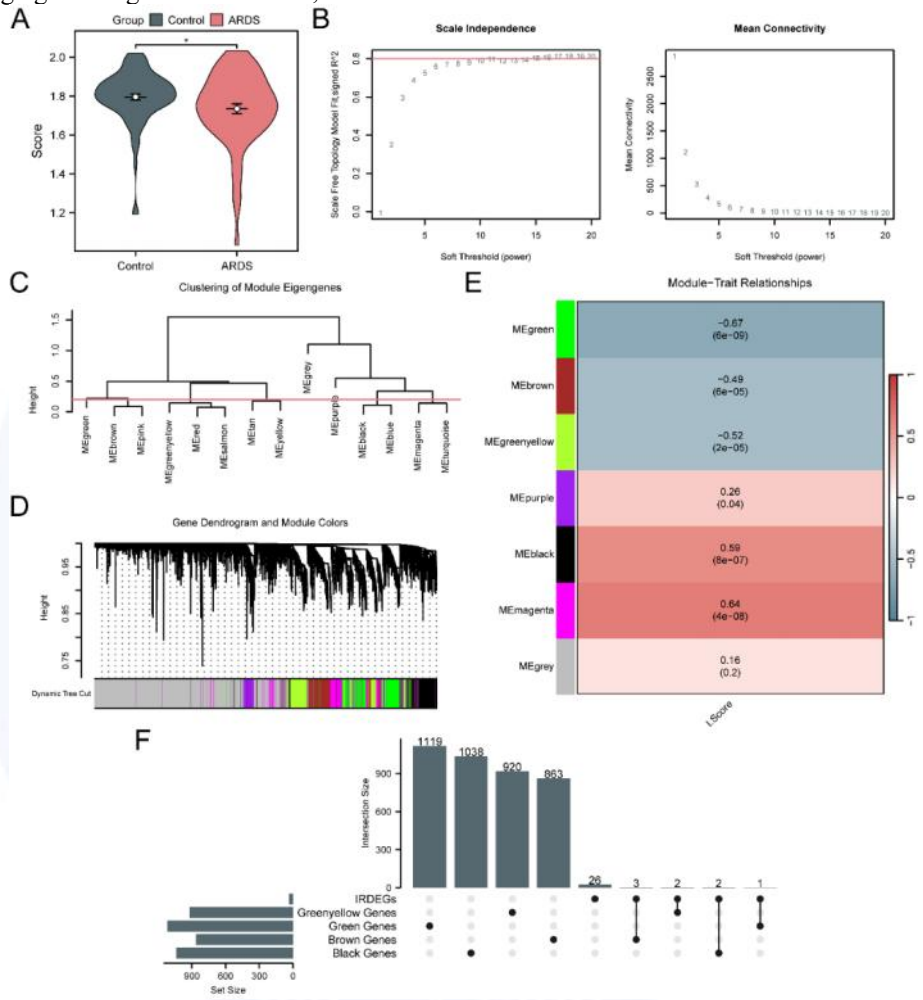


Note: A, volcano plot of DEGs; B, Venn diagram of the intersection between DEGs and IRGs; C, expression heatmap of IRDEGs; D, differential expression ranking plot of DEGs; E, bar plot of GO and KEGG pathway enrichment analysis of IRDEGs.

Fig.1 Identification and functional analysis results of IRDEGs

2.3 Construction of I.Score and WGCNA

Based on the expression profile of 34 IRDEGs, I.Score was calculated using ssGSEA, and there was a statistically significant difference in I.Score between the ARDS group and the control group (Figure 2A). WGCNA was performed on the top 70% genes with the highest expression variance in the ARDS group. The optimal soft threshold was determined to be 15 through scale-free fitting index, the minimum module gene number was set to 100, the module merging cut height was set to 0.2, and the minimum distance was set to 0.2. Finally, 7 co-expression modules were identified (Figure 2B-Figure 2D). Among them, 5 modules were significantly correlated with I.Score (|r| > 0.40) (Figure 2E). By intersecting the genes of these 5 modules with IRDEGs, 8 core module genes were finally screened out: TNFRSF1B, TNFSF10, B2M, S100 calcium-binding protein A9 (S100A9), SEMA7A, CD200, signal transducer and activator of transcription 3 (STAT3), and C-X-C motif chemokine ligand 12 (CXCL12) (Figure 2F).



Note: A, comparison of I.Score between the ARDS group and the control group, $^{*}P < 0.05$; B, optimal soft threshold screening and network topology analysis; C, clustering dendrogram of gene modules; D, module assignment results; E, correlation between each module and I.Score; F, intersection analysis of key module genes and IRDEGs (Upset plot).

Fig.2 WGCNA plot

2.4 Construction and validation of ARDS diagnostic model

Based on the 8 module genes, logistic regression screened out 6 genes with diagnostic value. Further through random forest and LASSO regression, a total of 4 model genes, namely B2M, CD200, SEMA7A and TNFSF10, were finally determined to construct the diagnostic model (Supplementary Figure 1, scan the QR code on the first page of the article). The nomogram showed that B2M contributed the most, while TNFSF10 contributed the least (Figure 3A). Calibration curve and decision curve analysis indicated that the model had consistent prediction performance and clinical utility (Figure 3B, 3C). The risk value was calculated according to the risk score formula (RiskScore = -B2M × 1.543 + CD200 × 1.388 + SEMA7A × 0.183 + TNFSF10 × 0.041), with a median of -9.6974. The ROC curve constructed using RiskScore showed good diagnostic efficacy (AUC = 0.745, Figure 3D).

2.5 GSEA of high- and low-risk groups

A total of 3 341 DEGs were identified between the high- and low-risk groups of ARDS, among which 1 827 were upregulated and 1 514 were downregulated

(Supplementary Figure 2A, 2B). GSEA analysis showed that these genes were significantly enriched in biological processes and pathways such as oxidative phosphorylation, respiratory electron transport, interleukin-12 signaling pathway, and FcεRI-mediated NF-κB activation (Supplementary Figure 2C-2G).

2.6 Immune infiltration analysis

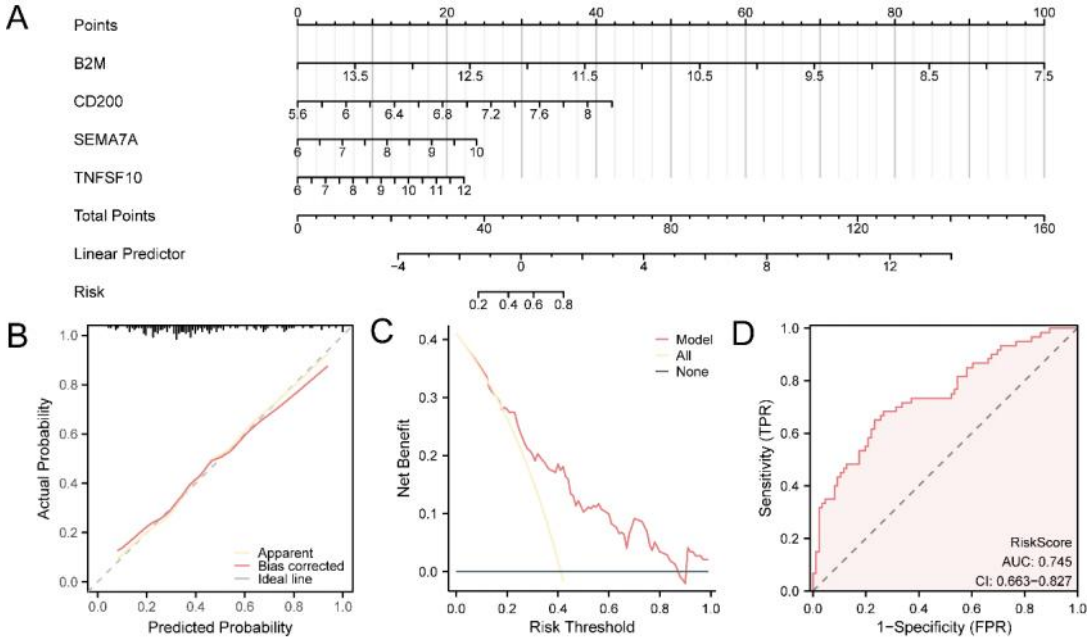
Immune infiltration analysis showed that 21 types of immune cells were significantly enriched in ARDS (Supplementary Figure 3A). Correlation analysis showed that follicular helper T cells had the strongest positive correlation with resting dendritic cells ($r = 0.765$, $P < 0.05$) (Supplementary Figure 3B). In the correlation analysis between model genes and immune cells, B2M had the strongest negative correlation with naive B cells ($r = -0.383$, $P < 0.05$) (Supplementary Figure 3C).

2.7 Immune infiltration analysis of high- and low-risk groups

Between the high- and low-risk groups of ARDS, a total of 8 types of immune cells showed statistically significant differences in infiltration abundance (Supplementary Figure 3D). Correlation analysis showed that central memory $CD4^+$ T cells were strongly positively correlated with myeloid-derived suppressor cells in both the low-risk group ($r = 0.713$, $P < 0.05$) and the high-risk group ($r = 0.865$, $P < 0.05$), with a stronger correlation in the latter (Supplementary Figure 3E, 3F). Correlation analysis between model genes and immune cells showed that in the low-risk group, SEMA7A had the strongest negative correlation with central memory $CD4^+$ T cells ($r = -0.623$, $P < 0.05$); in the high-risk group, TNFSF10 had the strongest positive correlation with neutrophils ($r = 0.784$, $P < 0.05$) (Supplementary Figure 3G, 3H).

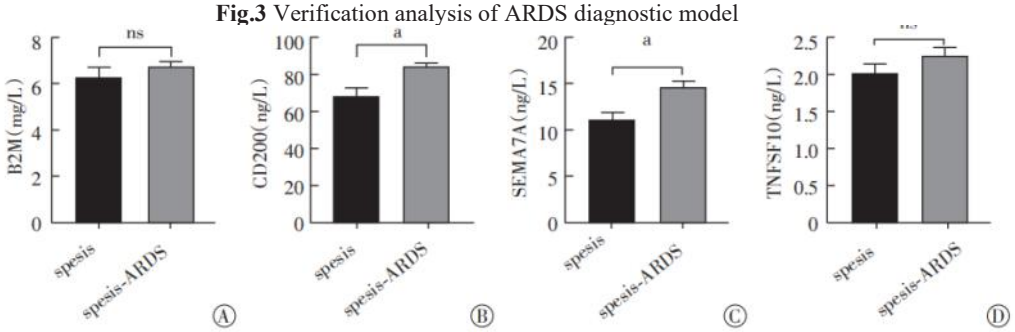
2.8 Serum levels of CD200, SEMA7A, B2M and TNFSF10 in clinical trial patients

The study of clinical trial samples showed that compared with patients in the Sepsis group, the serum levels of CD200 and SEMA7A in patients in the Sepsis-ARDS group were significantly increased ($P < 0.05$) (Figure 4).



Note: A, nomogram of the diagnostic model; B, calibration curve of the model; C, decision curve analysis plot (Y-axis: net benefit; X-axis: threshold probability); D, ROC curve of RiskScore in the comprehensive dataset.

Fig.3 Verification analysis of ARDS diagnostic model



Note: A, B2M; B, CD200; C, SEMA7A; D, TNFSF10; ns indicates $P > 0.05$; [a] indicates $P < 0.01$.

Fig.4 Comparison of serum marker levels between two groups

3. Discussion

ARDS is a disease characterized by excessive inflammation and immune imbalance, whose molecular mechanism is closely related to abnormalities in the immune regulatory network. Despite the in-depth understanding of this disease, the improvement in patient prognosis remains limited[1,13]. In previous studies, Yang *et al.*[14] identified 8 hub genes closely associated with type II alveolar epithelial cell injury; Liang et al.[15] confirmed that colony-stimulating factor 2 has potential diagnostic value and therapeutic potential in the lipopolysaccharide (LPS)-induced ALI model. Meanwhile, candidate biomarkers such as SIGLEC9, TSPO, CKS1B, and PTTG3P have shown promising application prospects in the early prediction of ARDS[16]. Research on immune-related biomarkers is of great significance for promoting the development of precision medicine for ARDS.

In this study, 34 IRDEGs were identified. Functional enrichment analysis showed that these genes are mainly involved in key processes such as pathogen recognition, cytokine signal transduction, and necroptosis. Further GO and KEGG analyses revealed that they may amplify inflammatory responses by localizing to the outer plasma membrane and membrane rafts and regulating signaling pathways such as NF-κB[17]. The significant enrichment of cytokine-cytokine receptor interaction and necroptosis-related pathways indicates that the receptor-interacting serine/threonine-protein kinase 1/mixed lineage kinase domain-like pseudokinase (RIPK1/MLKL) axis-mediated cell death pathway may exacerbate alveolar injury by releasing damage-associated molecular patterns[18-19]. In addition, the enrichment of rheumatoid arthritis-related pathways suggests that autoantibodies may attack components of the lung tissue basement membrane through molecular mimicry mechanisms and promote the progression of pulmonary fibrosis, which is consistent with the results of multiple studies[20-21]. Previous studies have shown that inhibiting NF-κB signaling and improving mitochondrial oxidative phosphorylation can promote endothelial barrier repair and prevent LPS-induced ARDS[22]. Lee *et al.*[23] also found that mitochondria isolated from human stem cells exert anti-inflammatory effects on alveolar macrophages. In this study, the co-enrichment of oxidative phosphorylation and NF-κB pathways indicates that mitochondrial dysfunction and persistent inflammatory response may synergistically drive disease progression, providing a basis for targeted therapy targeting the energy metabolism-inflammation axis.

In this study, a 4-gene diagnostic model consisting of *B2M*, *CD200*, *SEMA7A*, and *TNFSF10* was constructed through logistic regression, random forest, and LASSO regression, which showed good performance in identifying ARDS. Further verification revealed that serum levels of CD200 and SEMA7A were significantly elevated in patients with sepsis-induced ARDS. Compared with single biomarkers, this multi-gene model can more comprehensively reflect disease heterogeneity and has higher specificity in distinguishing sepsis-induced ARDS from simple sepsis. CD200 mainly exerts its effects by binding to the CD200 receptor (CD200R), regulating the activity of microglia and macrophages, and inhibiting immune cell activation and inflammatory factor release[24]. Patoine *et al.*[25] observed that pulmonary inflammation was exacerbated in CD200-deficient rat ARDS models. The CD200/CD200R pathway exerts a selective regulatory effect on acute pneumonia, suggesting that the addition of CD200 agonists to multi-target treatment strategies may bring potential benefits. The overexpression of CD200 may reflect the negative regulation of excessive inflammation by the body[24]. SEMA7A activates the NF-κB pathway through receptors such as integrin β1, promotes the release of pro-inflammatory cytokines, and exacerbates inflammation[26]. In addition, SEMA7A promotes neutrophil chemotaxis and aggregation while disrupting the pulmonary endothelial barrier, further aggravating disease progression[27]. The results of this study are consistent with previous mechanistic studies, providing a basis for multi-target diagnosis and treatment strategies.

In this study, the complex regulatory network of the immune microenvironment in sepsis-induced ARDS was revealed through immune infiltration analysis. Significant enrichment of 21 types of immune cells indicated widespread immune activation and recruitment. Resting dendritic cells, as a subset of immature dendritic cells, can transform into an activated state under inflammatory stimulation, recruiting and activating T cell subsets[28]. Follicular helper T cells drive B cells to differentiate into memory B cells or plasma cells in the germinal center, regulating antibody class switching and the production of high-affinity antibodies[29-30]. Previous studies have found that follicular helper T cells and resting dendritic cells are simultaneously enriched in a variety of tumor lesions and are associated with improved prognosis[31-32]. This study confirmed a strong positive correlation between follicular helper T cells and resting dendritic cells in ARDS samples (r = 0.765). We speculate that the enrichment of resting dendritic cells and follicular helper T cells may promote pathogen clearance and facilitate ARDS recovery. The analysis also showed a significant negative correlation between B2M and naive B cells in ARDS samples (r = -0.383), which confirms the immune homeostasis imbalance commonly observed in critical respiratory diseases. B2M is an essential subunit of MHC-I molecules, responsible for maintaining the stability and function of the MHC-I complex and is crucial for antigen presentation[33]. Its downregulated expression may lead to defects in local antigen presentation, affecting the helper effect of CD4+ T cells on B cells[34], thereby weakening antibody production and pathogen clearance capacity.

The infiltration of immune cells between the high- and low-risk ARDS groups showed high heterogeneity, among which the abundance differences of 8 types of immune cells were statistically significant. This study found that central memory CD4+ T cells and myeloid-derived suppressor cells were strongly positively correlated in both groups. The expansion of the latter in the acute phase of ARDS may reflect negative feedback regulation of excessive inflammation, but it can also inhibit

T cell function by secreting interleukin-10 or consuming arginine, thereby exacerbating immune imbalance[35-36]. In the high-risk group, the expression of the apoptosis-inducing ligand TNFSF10 was significantly positively correlated with neutrophil infiltration. Its elevation may be related to the resistance of neutrophils to apoptotic signals or impaired clearance, which reflects the pathological state of abnormal survival and sustained activation of neutrophils in ARDS[37]. In the low-risk group, SEMA7A showed the strongest negative correlation with central memory CD4+ T cells. Previous studies have found that SEMA7A not only promotes inflammatory responses but also indirectly inhibits T cell function by regulating other members of the Semaphorin family[38]. Elevated SEMA7A may indirectly affect the proportion of this T cell subset by activating immunosuppressive pathways, which may reflect the recovery trend of immune homeostasis in the low-risk group.

This study has several limitations: first, the specific mechanisms of the identified genes need to be further explored; second, larger-scale and more diverse studies are required to validate these findings.; finally, the integration of clinical data will enhance the objective evaluation of the results. In the future, we will further strengthen the collection of clinical information from patients and promote the translation of these findings into clinical practice.

In summary, this study identified 4 IRDEGs with diagnostic potential for ARDS, clarified the associations between model genes and immune cells as well as among different immune cells, and confirmed the differential expression of SEMA7A and CD200 between patients with sepsis-induced ARDS and patients with sepsis alone.

Conflict of interest None

Reference

[1] Bos LDJ, Ware LB. Acute respiratory distress syndrome: causes, pathophysiology, and phenotypes[J]. Lancet, 2022, 400(10358): 1145-1156.

[2] Bellani G, Laffey JG, Pham T, et al. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries[J]. JAMA, 2016, 315(8): 788-800.

[3] Huppert LA, Matthay MA, Ware LB. Pathogenesis of acute respiratory distress syndrome[J]. Semin Respir Crit Care Med, 2019, 40(1): 31-39.

[4] Aegerter H, Lambrecht BN, Jakubzick CV. Biology of lung macrophages in health and disease[J]. Immunity, 2022, 55(9): 1564-1580.

[5] Yu WJ, Yang ML, Lv BW, et al. CD40L-activated DC promotes Th17 differentiation and inhibits Th2 differentiation in sepsis-induced lung injury via cGAS-STING signaling[J]. Biochem Genet, 2025, 63(3): 2455-2469.

[6] Zhong XH, Jin JY, Zhang H, et al. Group 2 innate lymphoid cells derived IL-9 reduces macrophage apoptosis and attenuates acute lung injury in sepsis[J]. Inflamm Res, 2025, 74(1): 38.

[7] Cheng L, Jiao Y, Jiang W, et al. IL-33 deficiency attenuates lung inflammation by inducing Th17 response and impacting the Th17/treg balance in LPS-induced ARDS mice via dendritic cells[J]. J Immunol Res, 2022, 2022: 9543083.

[8] Guervilly C, Fournier T, Chommeloux J, et al. Ultra-lung-protective ventilation and biotrauma in severe ARDS patients on veno-venous extracorporeal membrane oxygenation: a randomized controlled study[J]. Crit Care, 2022, 26(1): 383.

[9] Sun MF, Yang QQ, Hu CL, et al. Identification and validation of autophagy-related genes in sepsis-induced acute respiratory distress syndrome and immune infiltration[J]. J Inflamm Res, 2022, 15: 2199-2212.

[10] Wu XL, Guo YN. Role of cellular senescence genes and immune infiltration in sepsis and sepsis-induced ARDS based on bioinformatics analysis[J]. J Inflamm Res, 2024, 17: 9119-9133.

[11] Liu TT, Gao L, Li XY. Identification of diagnostic biomarkers and therapeutic targets in sepsis-associated ARDS via combining bioinformatics with machine learning analysis[J]. J Inflamm Res, 2025, 18: 9523-9536.

[12] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update[J]. Nucleic Acids Res, 2013, 41(Database issue): D991-D995.

[13] Gorman EA, O'Kane CM, McAuley DF. Acute respiratory distress syndrome in adults: diagnosis, outcomes, long-term sequelae, and management[J]. Lancet, 2022, 400(10358): 1157-1170.

[14] Yang XT, Wang J, Liu W. Molecular markers of type II alveolar epithelial cells in acute lung injury by bioinformatics analysis[J]. Sci Rep, 2023, 13(1): 17797.

[15] Liang QC, Zhou Q, Li JH, et al. Validation of novel hub genes and molecular mechanisms in acute lung injury using an integrative bioinformatics approach[J]. Acta Biochim Biophys Sin, 2021, 53(3): 342-353.

[16] Ming TQ, Dong MY, Song XM, et al. Integrated analysis of gene co-expression network and prediction model indicates immune-related roles of the identified biomarkers in sepsis and sepsis-induced acute respiratory distress syndrome[J]. Front Immunol, 2022, 13: 897390.

[17] Wang KT, Yang JL, Deng JP, et al. Pinocembrin reduces pyroptosis to improve flap survival by modulating the TLR4/NF-κB/NLRP3 signaling pathway[J]. Biochim Biophys Acta Mol Basis Dis, 2025, 1871(3): 167710.

[18] Huang HR, Cho SJ, Harris RM, et al. RIPK3 activates MLKL-mediated necroptosis and inflammasome signaling during Streptococcus Infection[J]. Am J Respir Cell Mol Biol, 2021, 64(5): 579-591.

[19] Liu YL, Zhu H, Chen H, et al. LPS-induced TMBIM6 splicing drives endothelial necroptosis and aggravates ALI[J]. Respir Investig, 2025, 63(2): 191-199.

[20] Leavy OC, Kawano-Dourado L, Stewart ID, et al. Rheumatoid arthritis and idiopathic pulmonary fibrosis: a bidirectional Mendelian randomisation study[J]. Thorax, 2024, 79(6): 538-544.

[21] Zhao R, Zhang YW, Guo JC, et al. Genetic evidence reveals a causal relationship between rheumatoid arthritis and interstitial lung disease[J]. Front Genet, 2024, 15: 1395315.

[22] Peng JN, Tang R, He J, et al. S1PR3 inhibition protects against LPS-induced ARDS by inhibiting NF-κB and improving mitochondrial oxidative phosphorylation[J]. J Transl Med, 2024, 22(1): 535.

[23] Lee SE, Kim IH, Kang YC, et al. Mitochondrial transplantation attenuates lipopolysaccharide-induced acute respiratory distress syndrome[J]. BMC Pulm Med, 2024, 24(1): 477.

[24] Li DP, Wang Y, Tang L, et al. CD200-CD200R1 signalling attenuates imiquimod-induced psoriatic inflammation by inhibiting the activation of skin inflammatory macrophages[J]. Int Immunopharmacol, 2020, 78: 106046.

[25] Patoine D, Bouchard K, Lemay AM, et al. Specificity of CD200/CD200R pathway in LPS-induced lung inflammation[J]. Front Immunol, 2022, 13: 1092126.

[26] Li X, Xie WL, Pan Q, et al. Semaphorin 7A interacts with nuclear factor NF-kappa-B p105 via integrin β1 and mediates inflammation[J]. Cell Commun Signal, 2023, 21(1): 24.

[27] Granja T, Köhler D, Tang LY, et al. Semaphorin 7A coordinates neutrophil response during pulmonary inflammation and sepsis[J]. Blood Adv, 2024, 8(11): 2660-2674.

[28] Kubo M, Harada Y, Sasaki T. The role of dendritic cells in the instruction of helper T cells in the allergic March[J]. Int Immunol, 2024, 36(11): 559-566.

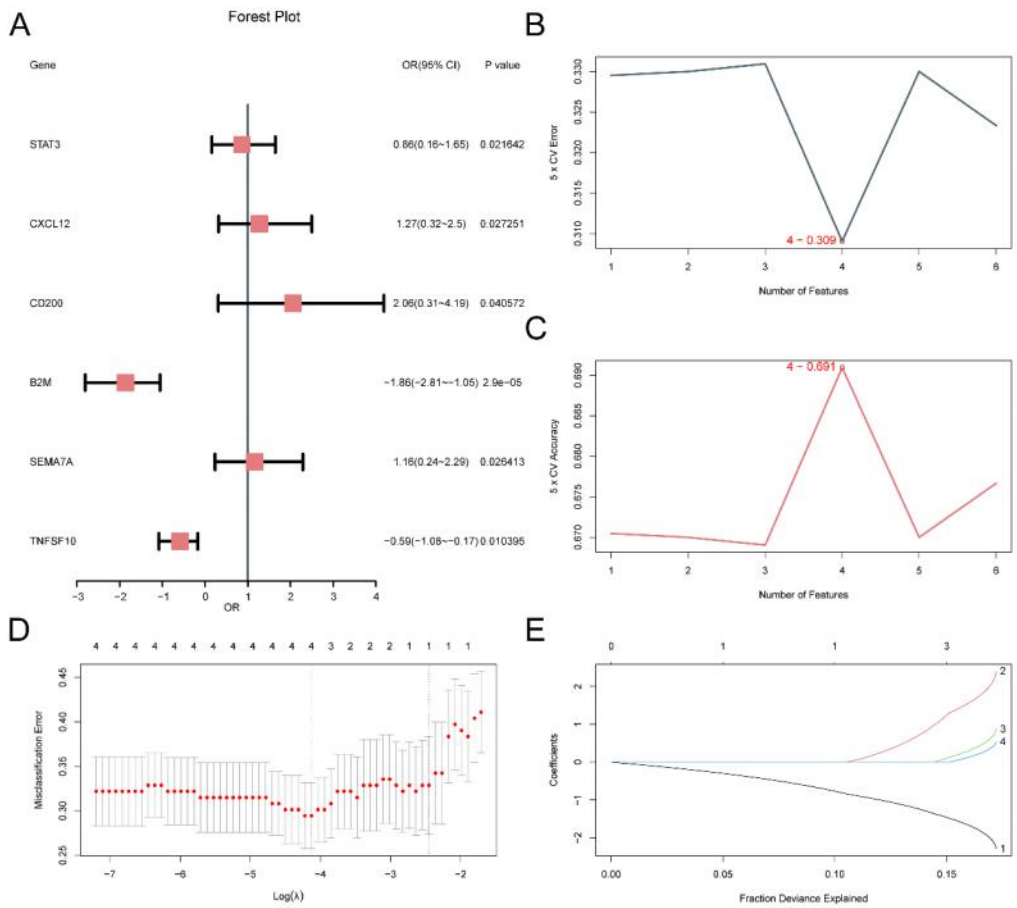
[29] Gulubova MV, Valkanov SP, Ignatova MMK, et al. Tertiary lymphoid structures in colorectal cancer - organization and immune cell interactions[J]. Am J Clin Exp Immunol, 2024, 13(6): 236-245.

[30] Phanthunane C, Wijers R, De Herdt MJ, et al. Intratumoral niches of B cells and follicular helper T cells, and the absence of regulatory T cells,

associate with longer survival in early-stage oral tongue cancer patients[J]. Cancers, 2022, 14(17): 4298.
[31] Li C, Yu S, Chen J, et al. Risk stratification based on DNA damage-repair-related signature reflects the microenvironmental feature, metabolic status and therapeutic response of breast cancer[J]. Front Immunol, 2023，14:1127982.
[32] Liu GH, Wang LP, Ji LL, et al. Identifying prognostic markers in spatially heterogeneous breast cancer microenvironment[J]. J Transl Med, 2023, 21(1): 580.
[33] Cruz FM, Orellano LAA, Chan A, et al. Alternate MHC I antigen presentation pathways allow CD8+ T-cell recognition and killing of cancer cells in the absence of β2M or TAP[J]. Cancer Immunol Res, 2025, 13(1): 98-108.
[34] Han XW, Zhang JY, Li WD, et al. The role of B2M in cancer immunotherapy resistance: function, resistance mechanism, and reversal strategies[J]. Front Immunol, 2025, 16: 1512509.
[35] Dar AA, Patil RS, Pradhan TN, et al. Myeloid-derived suppressor cells impede T cell functionality and promote Th17 differentiation in oral squamous cell carcinoma[J]. Cancer Immunol Immunother, 2020, 69(6): 1071-1086.
[36] Pan PY, Ma G, Weber KJ, et al. Immune stimulatory receptor CD40 is required for T-cell suppression and T regulatory cell activation mediated by myeloid-derived suppressor cells in cancer[J]. Cancer Res, 2010, 70(1): 99-108.
[37] Song C, Li HT, Mao Z, et al. Delayed neutrophil apoptosis may enhance NET formation in ARDS[J]. Respir Res, 2022, 23(1): 155.
[38] Huang JH, Zhao CM, Zhang SW. Semaphorin 7A promotes endothelial permeability and inflammation via plexin C1 and integrin β1 in Kawasaki disease[J]. BMC Pediatr, 2024, 24(1): 285.

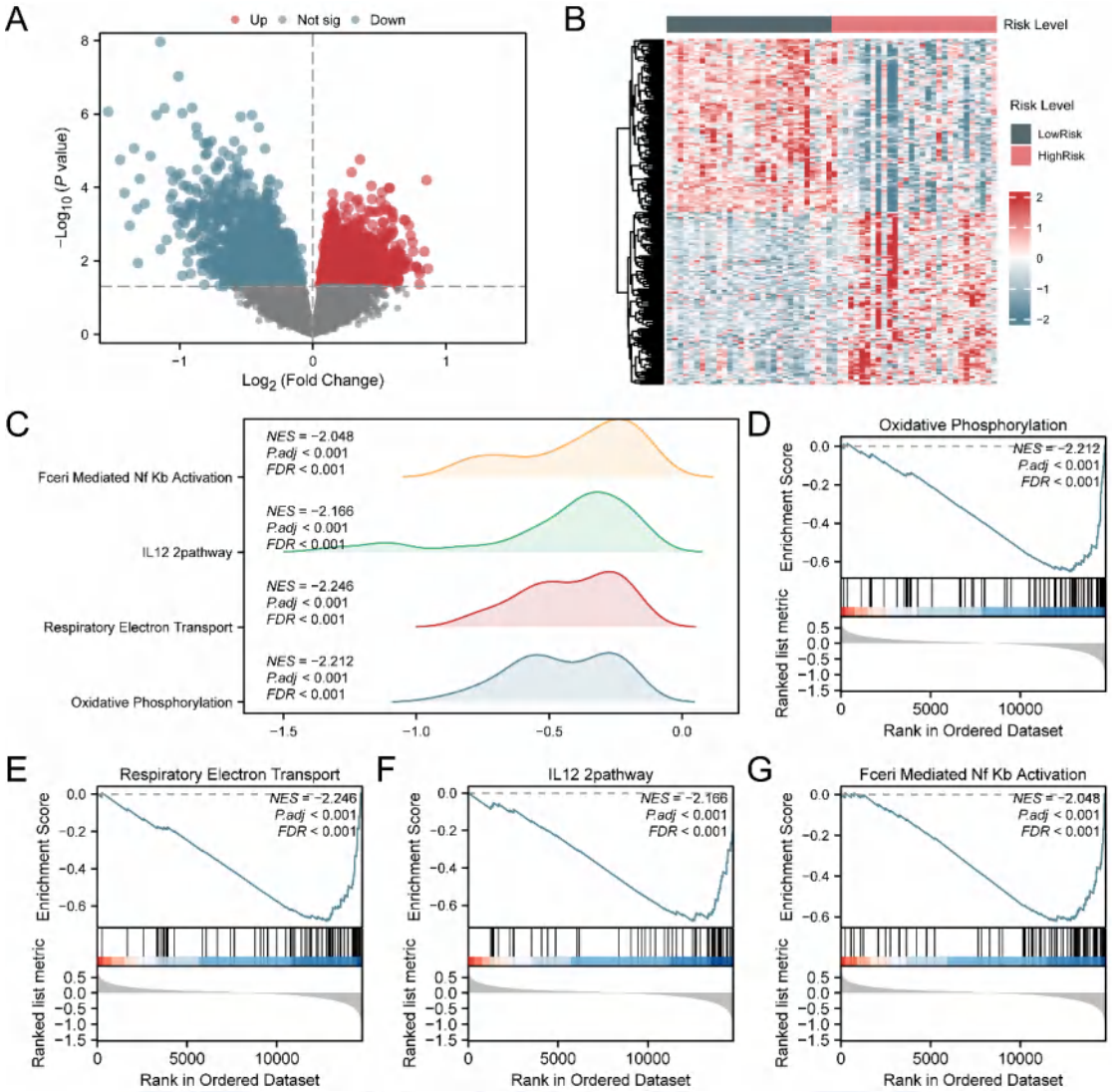
Submission Received: 2025-12-23/**Revised:**2026-03-10

Supplementary figures:



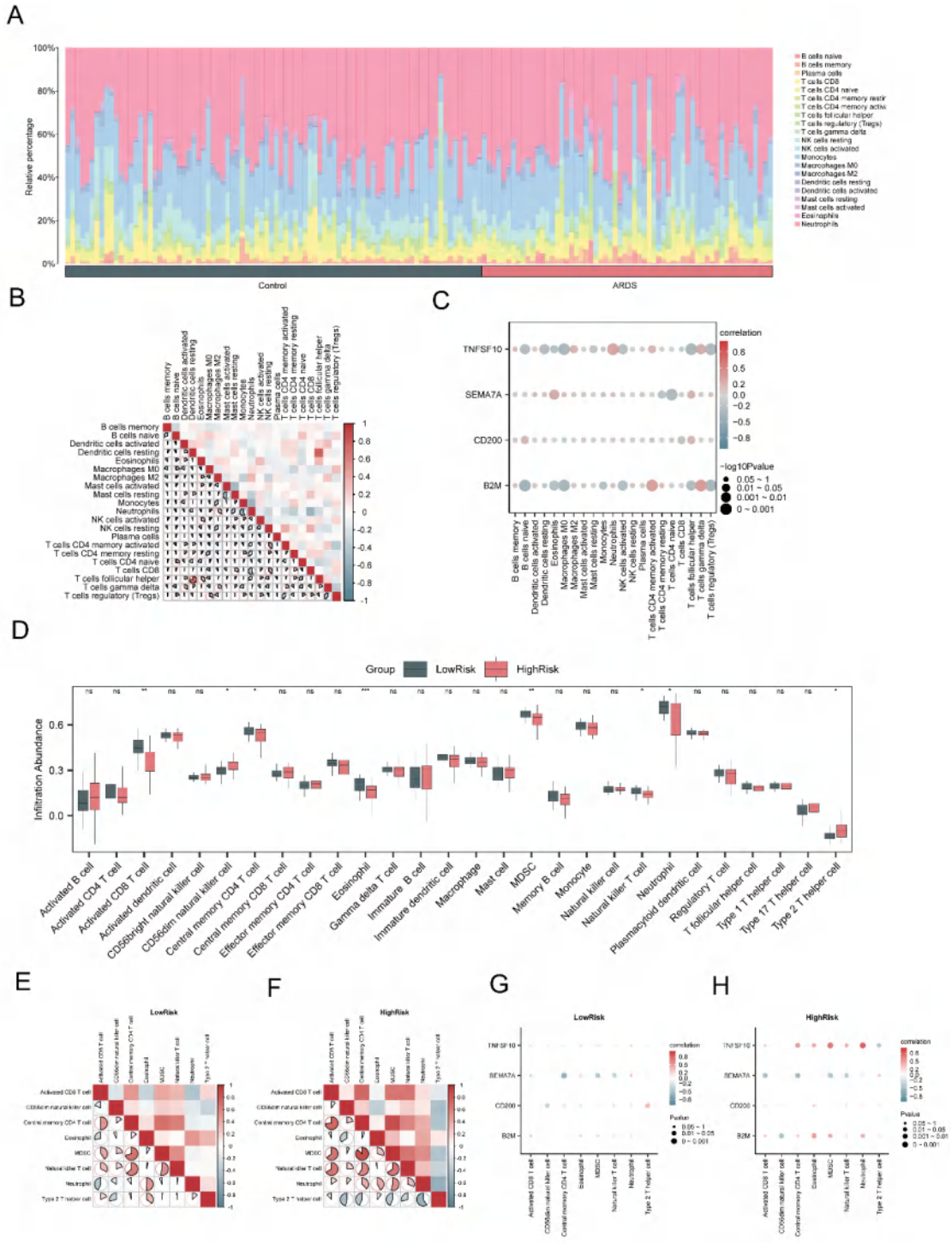
Note: A, Forest plot of module genes in logistic regression analysis; B-C,Optimal number of genes determined by the random forest algorithm based on error rate and accuracy; D-E, Diagnostic model plot and variable trajectory plot based on least absolute shrinkage and selection operator (LASSO) regression.

Supplementary Fig.1 Construction process of the ARDS diagnostic model



Note: A, volcano plot of differentially expressed genes; B, heatmap of differential gene expression; C, GSEA ridge plot of four biological functions; D-G, GSEA results showing significant enrichment of high ARDS risk in oxidative phosphorylation (D), respiratory electron transport (E), interleukin-12 (IL-12) signaling pathway (F), and Fc ε RI-mediated nuclear factor- κ B (NF- κ B) activation (G).

Supplementary Fig.2 Differential analysis and GSEA enrichment results between high- and low-risk groups



Note: A, immune cell proportion plot; B, correlation heatmap of immune cells; C, bubble plot showing the correlation between immune cell infiltration abundance and model genes; D, comparison of immune cells between the high- and low-risk groups; E-F, correlation heatmaps of immune cell infiltration abundance in the low-risk group (E) and high-risk group (F); G-H, bubble plots showing the correlation between immune cell infiltration abundance and model genes in the low-risk group (G) and high-risk group (H).

Supplementary Fig.3 Immune infiltration analysis results based on CIBERSORT and ssGSEA

·脓毒症专题·论著·

脓毒症诱导的急性呼吸窘迫综合征免疫相关生物标志物的鉴定和验证

熊雅欣[1]， 毛禹祁[1]， 谷全宽[2]， 赵鸣雁[1]， 孟祥林[1]
1. 哈尔滨医科大学附属第一医院重症医学科，黑龙江 哈尔滨 150000；
2. 通化市中心医院重症医学科，吉林 通化 134100

摘要：目的 通过识别免疫调节相关差异表达基因(IRDEGs)，为理解脓毒症诱导的急性呼吸窘迫综合征(ARDS)与免疫失衡相关的疾病机制及实现精准诊疗提供科学依据。**方法** 基于基因表达综合数据库(GEO)中的单纯脓毒症及脓毒症诱导的ARDS样本，筛选IRDEGs并进行基因本体(GO)和京都基因与基因组百科全书(KEGG)通路富集分析。通过加权基因共表达网络分析(WGCNA)获得共表达模块基因，结合logistic回归、随机森林和LASSO回归构建ARDS诊断模型。同时，利用CIBERSORT进行免疫浸润分析，并收集2024年就诊于哈尔滨医科大学附属第一医院的脓毒症诱导的ARDS与单纯脓毒症患者的外周血样本各10例，通过酶联免疫吸附试验(ELISA)验证模型基因的蛋白表达。**结果** 鉴定出34个IRDEGs和4个模型基因[肿瘤坏死因子配体超家族成员10(*TNFSF*10)、β2-微球蛋白(*B2M*)、信号素7A(*SEMA7A*)和分化簇200(*CD200*)]，并构建了有效的ARDS诊断模型。免疫浸润显示ARDS中免疫细胞广泛富集且相关性显著，如滤泡辅助T细胞与静息树突状细胞呈强正相关(r=0.765，P<0.05)。模型基因与免疫细胞浸润相关，如*B2M*与初始B细胞呈负相关(r=−0.383，P<0.05)。ELISA证实，与单纯脓毒症患者相比，脓毒症诱导的ARDS患者血清中CD200和SEMA7A水平显著升高(P<0.01)。**结论** 由*TNFSF10*、*B2M*、*SEMA7A*、*CD200*构建的ARDS诊断模型效能良好，模型基因和免疫细胞间以及各免疫细胞之间具有显著相关性，ELISA证实脓毒症ARDS患者血清CD200和SEMA7A显著升高，提示其为潜在诊断标志物。

关键词： 急性呼吸窘迫综合征；脓毒症；免疫调节；诊断模型
中图分类号： R631 R563.8 **文献标识码：** A **文章编号：** 1674-8182(2026)05-0669-08

Identification and validation of immune-related biomarkers for sepsis-induced acute respiratory distress syndrome

XIONG Yaxin*, MAO Yuqi, GU Quankuan, ZHAO Mingyan, MENG Xianglin
*Department of Critical Care Medicine, First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang 150000, China
Corresponding author: MENG Xianglin, E-mail: mengxianglin@hrbmu.edu.cn

Abstract: Objective To provide a scientific basis for understanding the disease mechanism and achieving precise diagnosis and treatment of sepsis-induced acute respiratory distress syndrome (ARDS) related to immune imbalance by identifying immune regulation-related differentially expressed genes (IRDEGs). **Methods** Based on sepsis alone and sepsis-induced ARDS samples gained from the Gene Expression Omnibus (GEO) database, IRDEGs were screened and subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. Weighted gene co-expression network analysis (WGCNA) was used to obtain co-expression module genes. Combined with logistic regression, random forest, and LASSO regression, an ARDS diagnostic model was constructed. Meanwhile, CIBERSORT was employed to analyze immune infiltration. Peripheral blood samples from patients with

DOI:10.13429/j.cnki.cjcr.2026.05.004
基金项目： 黑龙江省重点研发计划(GA21C011，JD22C005)
通信作者： 孟祥林，E-mail：mengxianglin@hrbmu.edu.cn
出版日期： 2026-05-20

QR code for English version

sepsis-induced ARDS and sepsis alone (n=10, each), treated at the First Affiliated Hospital of Harbin Medical University in 2024, were collected to validate the protein expression of model genes via enzyme-linked immunosorbent assay (ELISA). **Results** A total of 34 IRDEGs and 4 model genes [tumor necrosis factor superfamily member 10 (*TNFSF*10), β2-microglobulin (*B2M*), semaphorin 7A(*SEMA7A*), and cluster of differentiation 200(*CD200*)] were identified. An effective diagnostic model for ARDS was subsequently constructed. Immune infiltration analysis revealed extensive enrichment of immune cells and significant correlations in ARDS, such as a strong positive correlation between follicular helper T cells and resting dendritic cells (r=0.765, P<0.05). Model genes were associated with immune cell infiltration; for example, *B2M* was negatively correlated with naive B cells (r=−0.383, P<0.05). ELISA confirmed that serum levels of CD200 and SEMA7A were significantly higher in patients with sepsis-induced ARDS compared to those with sepsis alone (P<0.01). **Conclusion** The ARDS diagnostic model constructed using *TNFSF10*, *B2M*, *SEMA7A*, and *CD200* exhibit good performance. There are significant correlations between the model genes and immune cells, as well as among various immune cells themselves. ELISA confirms that serum levels of CD200 and SEMA7A are significantly elevated in patients with septic ARDS, suggesting that they are potential diagnostic biomarkers.

Keywords: Acute respiratory distress syndrome; Sepsis; Immune regulation; Diagnostic model

Fund program: Key Research and Development Plan of Heilongjiang Province (GA21C011, JD22C005)

急性呼吸窘迫综合征(acute respiratory distress syndrome, ARDS)是由多种病因引起的急性弥漫性肺损伤,可导致严重低氧血症[1]。脓毒症是其常见诱因,脓毒症诱导的ARDS具有更高的发病率和死亡率[2]。

在脓毒症诱导的ARDS病理进程中,免疫细胞持续激活并驱动炎症级联反应,释放炎症介质,导致外周免疫细胞浸润肺实质,破坏肺泡-毛细血管屏障,形成弥漫性肺水肿[3]。免疫调节涉及多种免疫细胞亚群及分子变化,是ARDS研究的重要方向。研究表明,急性肺损伤(acute lung injury, ALI)期间巨噬细胞的M1/M2表型极化平衡调控炎症进程[4],中性粒细胞、固有淋巴细胞和树突状细胞也在脓毒症诱导的ARDS发生发展中起重要作用[5-6]。临床研究证实白细胞介素-6、白细胞介素-8及肺表面活性蛋白D等生物标志物对ARDS患者预后有一定的预测价值[7-8]。目前该疾病的免疫分子网络尚未完全阐明,识别新型生物标志物对优化诊疗策略具有重要意义。近年来,基因表达特征已成为疾病研究的重点。Sun等[9]鉴定出7个自噬相关基因,能有效区分脓毒症诱导的ARDS和单纯脓毒症。另一项研究中,4个衰老相关基因被发现对脓毒症诱导的ARDS具有显著诊断价值[10]。此外,ARDS患者中性粒细胞相关基因表达也显著升高[11]。这些与特定生物过程相关的基因特征,有望作为脓毒症诱导的ARDS的新型诊断标志物及机制研究工具。

本研究通过生物信息学分析鉴定ARDS免疫调节相关差异表达基因(immune regulation-related differentially expressed genes, IRDEGs),并基于相关基因[肿瘤坏死因子配体超家族成员10(tumor necrosis factor ligand superfamily member 10, *TNFSF*10)、β2-微球蛋白(beta-2-microglobulin, *B2M*)、信号素7A(semaphorin 7A, *SEMA7A*)和分化簇200(cluster of differentiation 200, *CD200*)]进行ARDS诊断模型的构建。同时,通过利用估计RNA转录本相对子集进行细胞类型鉴定(cell-type identification by estimating relative subsets of RNA transcripts, CIBERSORT)算法分析脓毒症诱导的ARDS与单纯脓毒症间的免疫浸润差异。通过酶联免疫吸附试验(enzyme-linked immunosorbent assay, ELISA)进一步验证上述基因在两组患者中的表达水平,以期为理解疾病机制及实现ARDS的精准诊疗提供科学依据。

1　资料与方法

1.1　*数据源*　从基因表达综合数据库(Gene Expression Omnibus, GEO)[12]中获取数据集GSE32707(31例脓毒症诱导的ARDS及58例单纯脓毒症样本)和GSE66890(29例脓毒症诱导的ARDS及28例单纯脓毒症样本)。使用R包sva进行去批次处理后合并,获得包含60例ARDS和86例对照的综合数据集。同时,从GeneCards数据库筛选出237个免疫调节相关基因(immune regulation-related genes, IRGs),相关评分>1。

1.2　ARDS的IRDEGs　对两组样本进行差异分析,以筛选差异表达基因(differentially expressed genes, DEGs),基于差异倍数(fold change, FC),设定

$|\log_2 FC|>0.5$ 且 $P<0.05$ 为阈值。将DEGs与IRGs取交集，获得IRDEGs。随后，采用Spearman法分析这些基因表达量的相关性。

1.3 功能富集分析 使用clusterProfiler包对IRDEGs进行基因本体（Gene Ontology，GO）与京都基因与基因组百科全书（Kyoto Encyclopedia of Genes and Genomes，KEGG）通路富集分析。筛选标准为调整后 P 值<0.05且错误发现率（false discovery rate，FDR）<0.25，即 q 值<0.25。P 值校正方法为Benjamini-Hochberg（BH）。

1.4 免疫调节评分（I.Score）的构建和加权基因共表达网络分析（weighted gene co-expression network analysis，WGCNA） 使用单样本基因集富集分析（single-sample gene-set enrichment analysis，ssGSEA）算法，基于IRDEGs和综合数据集的表达矩阵计算所有样本的I.Score。对ARDS组中表达方差前70%基因进行WGCNA，获取共表达模块。筛选与I.Score显著相关的模块（$|r|>0.40$），并将其内基因与IRDEGs取交集，得到模块基因。

1.5 ARDS诊断模型的构建与验证 综合应用logistic回归、支持向量机及最小绝对收缩和选择算子（least absolute shrinkage and selection operator，LASSO）回归对模块基因进行筛选，获取关键模型基因以构建ARDS诊断模型。通过列线图评估基因权重，并结合校正曲线与决策曲线分析，分别评估模型校准度与临床净获益。基于LASSO风险系数计算风险评分（RiskScore），并通过受试者工作特征（receiver operating characteristic，ROC）曲线及曲线下面积（area under the curve，AUC）评估RiskScore的诊断效能。

1.6 高低风险组的基因集富集分析（gene-set enrichment analysis，GSEA） 基于RiskScore中位数将ARDS样本分为高风险组和低风险组，进行差异基因分析。从MSigDB数据库获取经典通路基因集，对两组样本进行GSEA富集分析。筛选标准为调整后 P 值<0.05且FDR（q 值）<0.25，P 值校正采用BH法。

1.7 免疫浸润分析 使用CIBERSORT算法结合白细胞基因特征矩阵22（leukocyte gene signature matrix 22，LM22）获得综合数据集的免疫细胞浸润谱。随后通过Spearman相关性分析，评估免疫细胞类型之间及其与模型基因表达的相关性。

1.8 高低风险组的免疫浸润分析 使用ssGSEA基于免疫细胞亚型计算浸润分数，构建免疫浸润矩阵。分析高低风险组间免疫细胞的差异表达，并基于Spearman相关分析免疫细胞间及其与模型基因的相关性。

1.9 临床试验验证 收集2024年就诊于哈尔滨医科大学附属第一医院的脓毒症诱导的ARDS（sepsis-ARDS组）与单纯脓毒症（sepsis组）患者的外周血样本各10例。于诊断脓毒症24 h内采集血样，经抗凝、离心后取血清，使用ELISA试剂盒定量检测TNFSF10、B2M、SEMA7A和CD200水平。本研究遵循《赫尔辛基宣言》，获哈尔滨医科大学附属第一医院伦理委员会批准（编号：IRB-AF/SC-04/0.2-0），所有受试者均签署知情同意书。

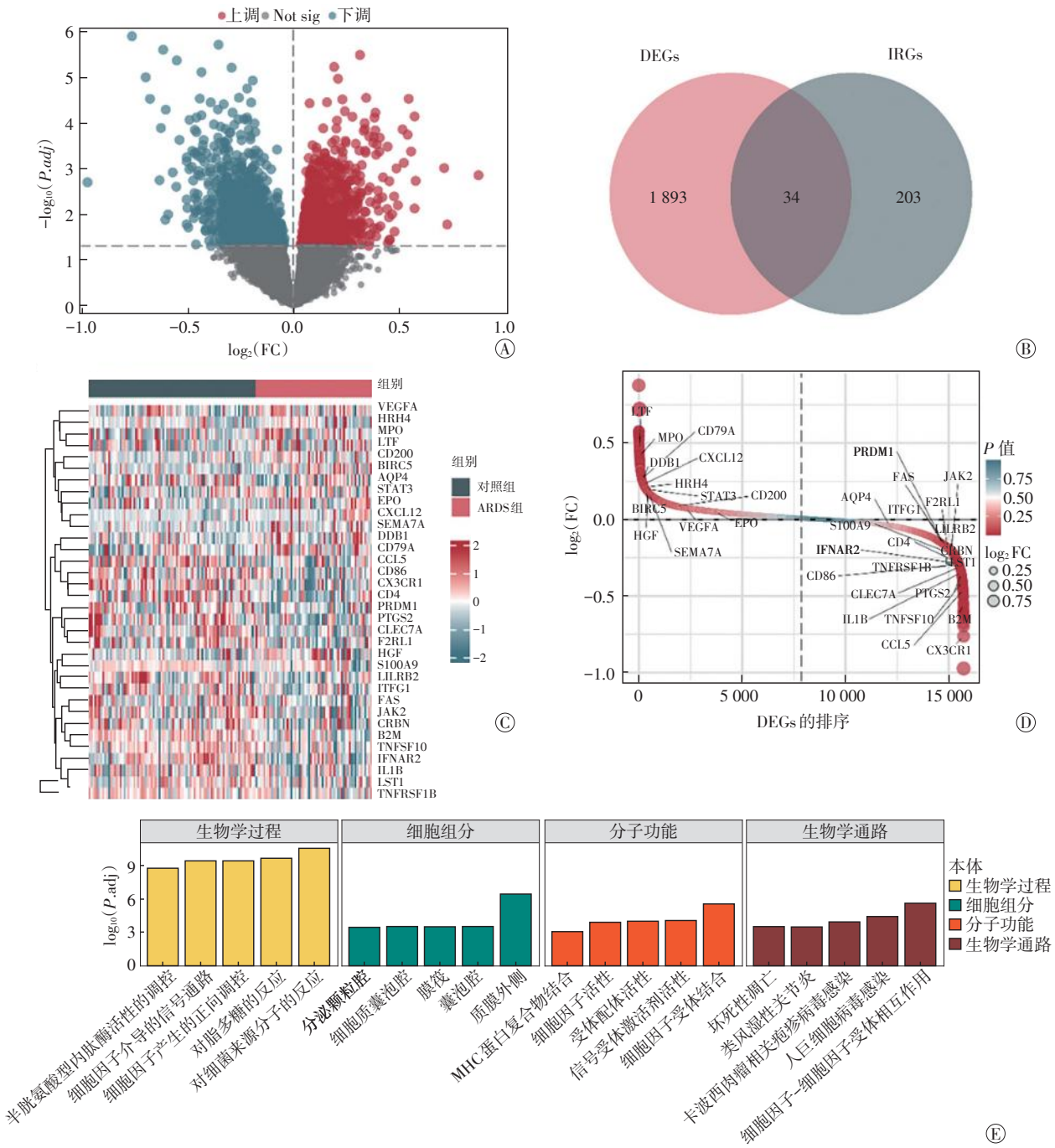
1.10 统计学方法 采用R 4.2.3软件进行统计分析及作图。两组间比较采用独立样本 t 检验，多组间比较采用单因素方差分析。

2 结 果

2.1 ARDS IRDEGs 在ARDS组和对照组样本中共鉴定出1 927个DEGs，其中1 021个上调，906个下调（图1A）。与IRGs取交集后，获得34个IRDEGs（图1B）。这些基因在两组间的表达差异通过热图及排序差异图进行可视化展示（图1C、图1D）。

2.2 功能富集分析 通过对IRDEGs进行GO和KEGG富集分析，结果显示其生物学过程富集于：对细菌来源分子的反应、对脂多糖的反应、细胞因子产生的正向调控、细胞因子介导的信号通路、半胱氨酸型内肽酶活性的调控等。细胞组分方面富集于：质膜外侧、分泌颗粒腔、细胞质囊泡腔、膜筏、囊泡腔等。分子功能方面富集于：细胞因子受体结合、受体配体活性、信号受体激活活性、细胞因子活性、主要组织相容性复合体（major histocompatibility complex，MHC）蛋白复合物结合等。同时，这些基因还富集在以下KEGG生物学通路中：细胞因子-细胞因子受体相互作用、人巨细胞病毒感染、卡波西肉瘤相关疱疹病毒感染、坏死性凋亡、类风湿性关节炎等（图1E）。

2.3 I.Score的构建和WGCNA 基于34个IRDEGs的表达谱，采用ssGSEA计算得到I.Score，其在ARDS组与对照组间差异有统计学意义（图2A）。对ARDS组中表达方差前70%的基因进行WGCNA分析，通过无标度拟合指数确定最优软阈值为15，设置最小模块基因数为100，模块合并剪切高度设为0.2，最小距离为0.2，最后识别出7个共表达模块（图2B~图2D）。其中5个模块与I.Score显著相关（$|r|>0.40$）（图2E）。将这5个模块的基因与IRDEGs取交集，最终筛选出8个核心模块基因：*TNFRSF1B*、*TNFSF10*、*B2M*、S100钙结合蛋白A9（S100 calcium-binding protein A9，



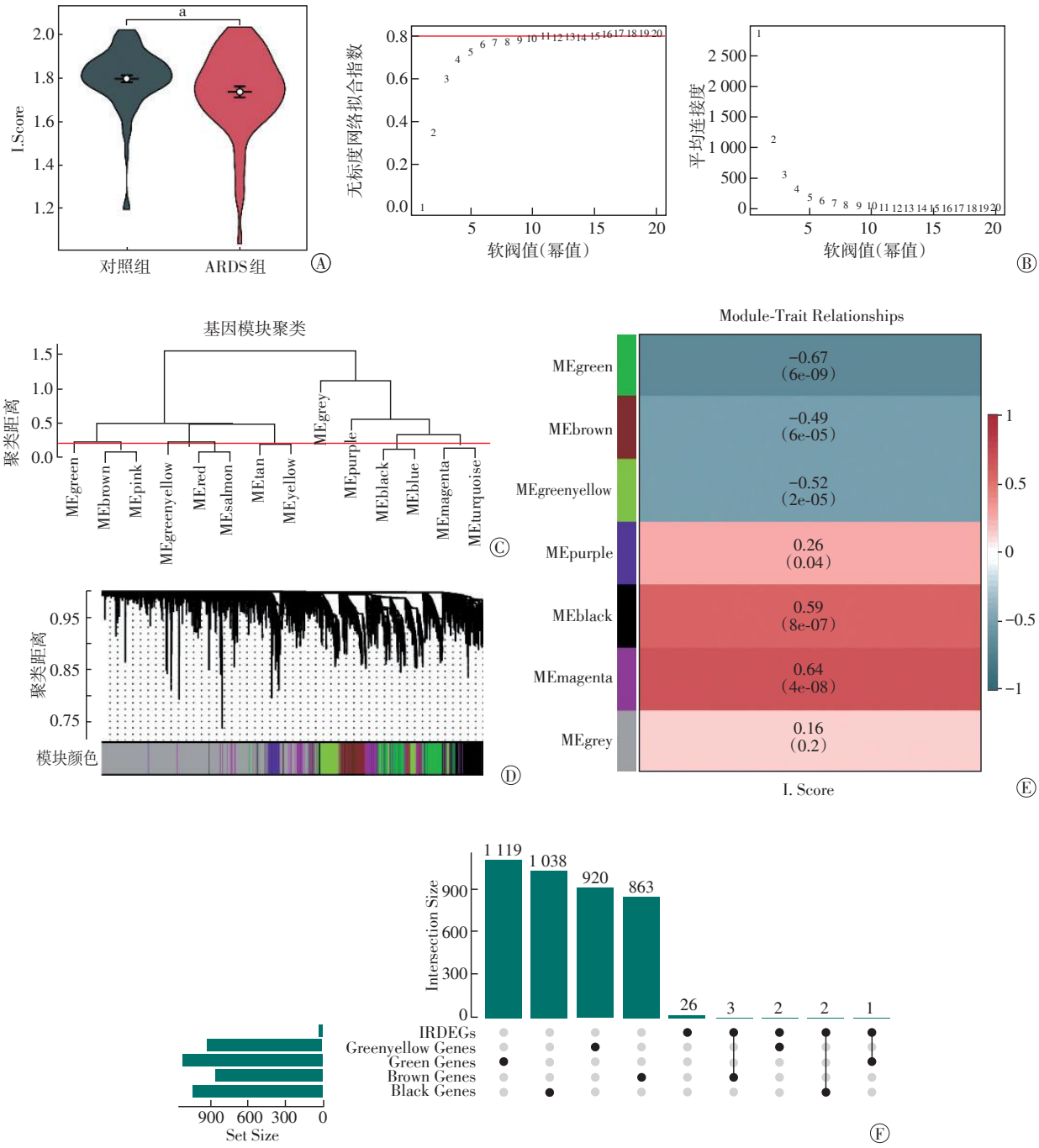
注：A，DEGs 的火山图；B，DEGs 与 IRGs 取交集的维恩图；C，IRDEGs 表达热图；D，DEGs 的差异排序图；E，IRDEGs 的 GO 与 KEGG 通路富集分析柱状图。

图 1　IRDEGs 的鉴定与功能分析结果

Fig.1　Identification and functional analysis results of IRDEGs

S100A9）、*SEMA7A*、*CD200*、信号转导与转录激活因子 3（signal transducer and activator of transcription，*STAT3*）和 C-X-C 基序趋化因子配体 12（C-X-C motif chemokine ligand，*CXCL12*）（图 2F）。

2.4　ARDS 诊断模型的构建与验证　基于 8 个模块基因，logistic 回归筛选出 6 个对诊断具有价值的基因。进一步通过随机森林和 LASSO 回归，最终确定 *B2M*、*CD200*、*SEMA7A* 和 *TNFSF10* 共 4 个模型基因构建诊断模型（见附加图 1，扫描文章首页二维码）。列线图显示 *B2M* 贡献最大，*TNFSF10* 最小（见图 3A）。校准曲线及决策曲线分析表明模型预测一致且具有临床实用性（见图 3B、3C）。根据风险评分公式（RiskScore=−B2M×1.543＋CD200×1.388＋SEMA7A ×0.183＋TNFSF10×0.041）计算风险值，其中位数为 −9.694。采用 Risk



注：A，ARDS 组与对照组的 I.Score 比较，^{a}P<0.05；B，最佳软阈值筛选及网络拓扑分析；C，基因模块聚类树状图；D，分配结果；E，各模块与 I.Score 的相关性；F，关键模块基因与 IRDEGs 的交集分析（Upset 图）。

图 2　WGCNA 图

Fig.2　WGCNA plot

Score 构建的 ROC 曲线显示其诊断效能良好（AUC=0.745，见图 3D）。

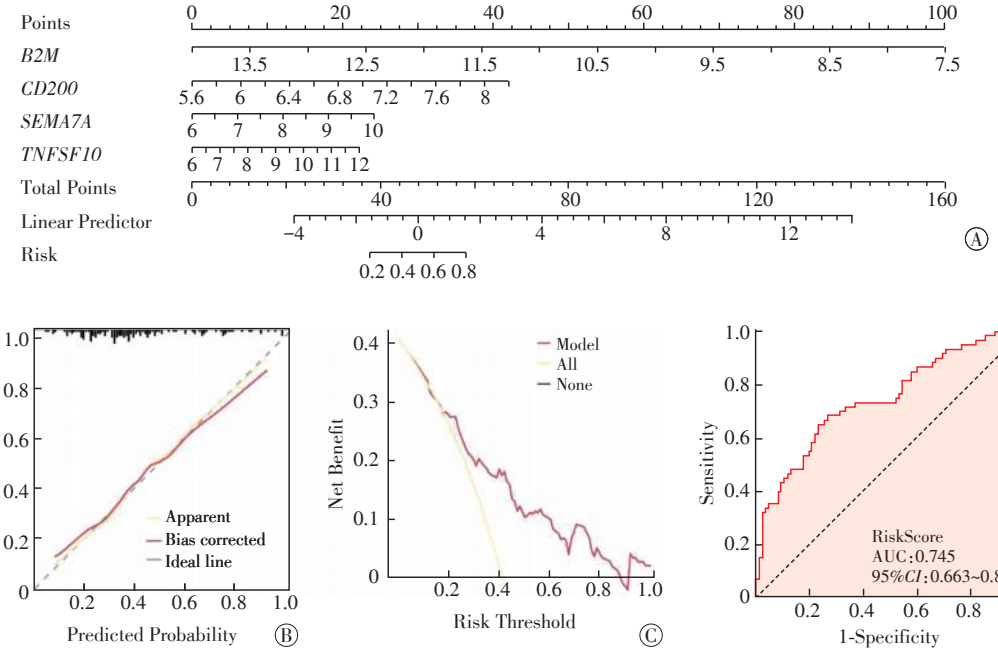
2.5　高低风险组的 GSEA　在 ARDS 的高、低风险组间共鉴定出 3 341 个 DEGs，其中上调 1 827 个、下调 1 514 个（见附加图 2A、2B）。GSEA 分析显示，这些基因显著富集于氧化磷酸化、呼吸链电子传递、白细胞介素-12 信号通路及 FcεRI 介导的 NF-κB 激活等生物学过程与通路（见附加图 2C~2G）。

2.6　免疫浸润分析　免疫浸润分析显示 21 种免疫细胞在 ARDS 中显著富集（见附加图 3A）。相关性分析表明，滤泡辅助性 T 细胞与静息树突状细胞呈最强正相关（r=0.765，P<0.05）（见附加图 3B）。模型基因与免疫细胞的相关性分析中，B2M 与初始 B 细胞呈最强负相关（r=−0.383，P< 0.05）（见附加

图 3C）。

2.7　高低风险组的免疫浸润分析　在 ARDS 高、低风险组间，共发现 8 种免疫细胞的浸润丰度差异有统计学意义（见附加图 3D）。相关性分析显示，中央记忆 CD4⁺T 细胞与髓源性抑制细胞在低风险组（r=0.713，P<0.05）及高风险组（r=0.865，P<0.05）均呈强正相关，且后者相关性更强（见附加图 3E、3F）。模型基因与免疫细胞的相关性分析表明，在低风险组中，*SEMA7A* 与中央记忆 CD4⁺T 细胞呈最强负相关（r=−0.623，P<0.05）；在高风险组中，*TNFSF10* 与中性粒细胞呈最强正相关（r=0.784，P<0.05）（见附加图 3G、3H）。

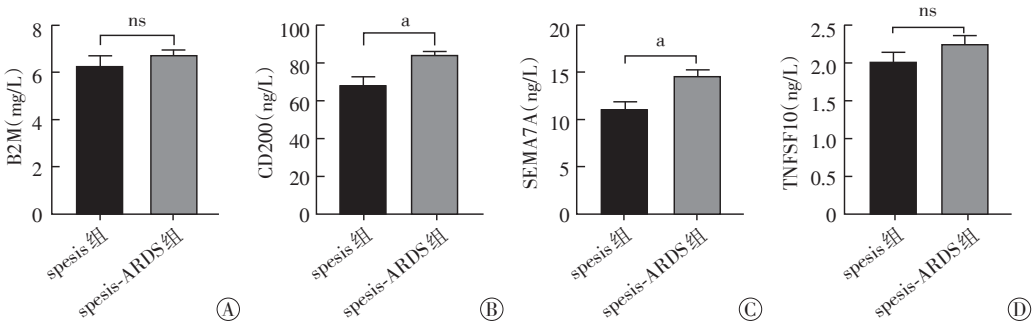
2.8　临床试验患者的 CD200、SEMA7A、B2M 和 TNFSF10 血清水平　对临床试验样本的研究显示，与 sepsis 组患者相比，sepsis-ARDS 组患者血清中 CD200 和 SEMA7A 的水平显著升高（P<0.05）。见图 4。



注：A，诊断模型的列线图；B，模型校准曲线；C，决策曲线分析图（Y 轴：净收益；X 轴：概率阈值）；D，RiskScore 在综合数据集中的 ROC 曲线。

图 3　ARDS 诊断模型的验证分析

Fig.3　Verification analysis of ARDS diagnostic model



注：A，B2M；B，CD200；C，SEMA7A；D，TNFSF10；ns 表示 P>0.05；a 表示 P<0.01。

图 4　两组临床患者血清标志物水平比较

Fig.4　Comparison of serum marker levels between two groups of clinical patients

3　讨　论

ARDS 作为一种以过度炎症和免疫失衡为特征的疾病，其分子机制与免疫调控网络异常密切相关。尽管对其认识有所深入，但患者预后改善有限[1,13]。先前研究中，Yang 等[14]鉴定出 8 个与Ⅱ型肺泡上皮细胞损伤密切相关的枢纽基因，Liang 等[15]证实集落刺激因子 2 在脂多糖（lipopolysaccharide，LPS）诱导的 ALI 模型中具有潜在诊断价值和治疗潜力，同时 SIGLEC9、TSPO、CKS1B 和 PTTG3P 等生物标志物候选物在早期 ARDS 预测方面展现出良好应用前景[16]。免疫相关生物标志物的研究对于推动 ARDS 精准医学发展意义重大。

本研究鉴定出 34 个 IRDEGs，其功能富集分析显

示这些基因主要参与病原体识别、细胞因子信号传导及坏死性凋亡等关键过程。GO与KEGG分析进一步揭示，它们可能通过质膜外侧及膜筏定位，调控NF-κB等信号通路放大炎症反应[17]，并且细胞因子-细胞因子受体相互作用和坏死性凋亡相关通路显著富集，表明RIPK1/MLKL轴介导的细胞死亡途径可能通过释放损伤相关分子模式加剧肺泡损伤[18-19]。此外，类风湿性关节炎相关通路的富集提示自身抗体可能通过分子模拟机制攻击肺组织基底膜成分，推动肺纤维化进程，这与多项研究的结果一致[20-21]。既往研究表明，抑制NF-κB信号传导和改善线粒体氧化磷酸化可促进内皮屏障修复，预防LPS诱导的ARDS[22]。Lee等[23]还发现，从人类干细胞中分离的线粒体对肺泡巨噬细胞有抗炎作用。本研究中，氧化磷酸化与NF-κB通路的共富集表明线粒体功能障碍与持续炎症反应可能协同驱动疾病进展，为针对能量代谢-炎症轴的靶向治疗提供了依据。

本研究通过logistic回归、随机森林和LASSO回归构建了由*B2M*、*CD200*、*SEMA7A*、*TNFSF10*组成的4基因诊断模型，该模型在鉴别ARDS方面表现良好。进一步验证发现，脓毒症诱导的ARDS患者血清中CD200与SEMA7A显著升高。相比单一生物标志物，该多基因模型能更全面反映疾病异质性，在区分脓毒症诱导的ARDS与单纯脓毒症时具有更高特异性。CD200主要通过结合CD200受体（CD200R）发挥作用，调节小胶质细胞和巨噬细胞的活性，抑制免疫细胞活化及炎症因子释放[24]。Patoine等[25]观察到，在*CD200*缺陷大鼠ARDS模型中肺部炎症加剧。CD200/CD200R通路对急性肺炎具有选择性调控作用，提示在多靶点治疗策略中加入CD200激动剂可能带来潜在获益。CD200的过表达可能反映机体对过度炎症的负向调节[24]。SEMA7A则通过整合素β1等受体激活NF-κB通路，促进促炎细胞因子释放并加剧炎症[26]。此外，SEMA7A促进中性粒细胞趋化和聚集，同时破坏肺内皮屏障，进一步加重疾病进展[27]。本研究结果与既往机制研究相符，为多靶点诊疗策略提供了依据。

本研究通过免疫浸润分析揭示了脓毒症诱导的ARDS中免疫微环境的复杂调控网络。21种免疫细胞显著富集，表明存在广泛的免疫激活与招募。静息树突状细胞作为一种未成熟树突状细胞亚群，在炎症刺激下可转化为激活状态，招募并激活T细胞亚群[28]。滤泡辅助T细胞在生发中心内驱动B细胞分化为记忆B细胞或浆细胞，调控抗体类别转换和高亲和力抗体的产生[29-30]。既往研究发现，在多种肿瘤病灶中，滤泡辅助T细胞和静息树突状细胞同时富集并与预后改善相关[31-32]。本研究证实ARDS样本中滤泡辅助T细胞和静息树突状细胞之间存在强正相关性（r=0.765），笔者推测静息树突状细胞和滤泡辅助T细胞的富集可能促进病原体清除并促进ARDS恢复。此分析还显示ARDS样本中B2M与初始B细胞之间存在显著负相关性（r=−0.383），印证了危重呼吸系统疾病中常见的免疫稳态失衡。B2M是MHC-I类分子的必需亚基，负责维持MHC-I复合体的稳定性和功能，对抗原呈递至关重要[33]。其表达下调可能导致局部抗原呈递缺陷，影响$CD4^+$T细胞对B细胞的辅助作用[34]，从而削弱抗体产生与病原体清除能力。

ARDS高、低风险组间的免疫细胞浸润呈现高度异质性，其中8种免疫细胞的丰度差异有统计学意义。本研究发现，两组样本的中央记忆$CD4^+$T细胞与髓源性抑制细胞均呈强正相关。后者在ARDS急性期的扩增可能反映对过度炎症的负反馈调节，但其也可通过分泌白细胞介素-10或消耗精氨酸抑制T细胞功能，从而加剧免疫失衡[35-36]。在高风险组中，凋亡诱导配体TNFSF10表达与中性粒细胞浸润呈显著正相关，其升高可能与中性粒细胞对凋亡信号产生抵抗或清除障碍有关，这反映了ARDS中中性粒细胞异常存活、持续活化的病理状态[37]。低风险组中，SEMA7A与中央记忆$CD4^+$T细胞呈最强负相关。既往研究发现，SEMA7A既促进炎症反应，又通过调节其他Semaphorin家族成员间接抑制T细胞功能[38]。SEMA7A升高可能通过激活免疫抑制途径间接影响该T细胞亚群的比例，这或许反映了低风险组中免疫稳态的恢复趋势。

本研究存在若干局限性：首先，已识别基因的具体机制需进一步探究。其次，需要更大规模和更多样化的研究来验证这些发现。最后，综合临床数据将增强对结果的客观评估。未来，笔者将进一步加强患者临床信息的获取，并促进其向临床实践转化。

综上所述，本研究识别了4个具有ARDS诊断潜力的IRDEGs，阐明了模型基因和免疫细胞间以及各免疫细胞之前的关联，并通过临床试验确认了脓毒症诱导的ARDS患者与脓毒症患者之间*SEMA7A*和*CD200*的差异表达。

利益冲突 无

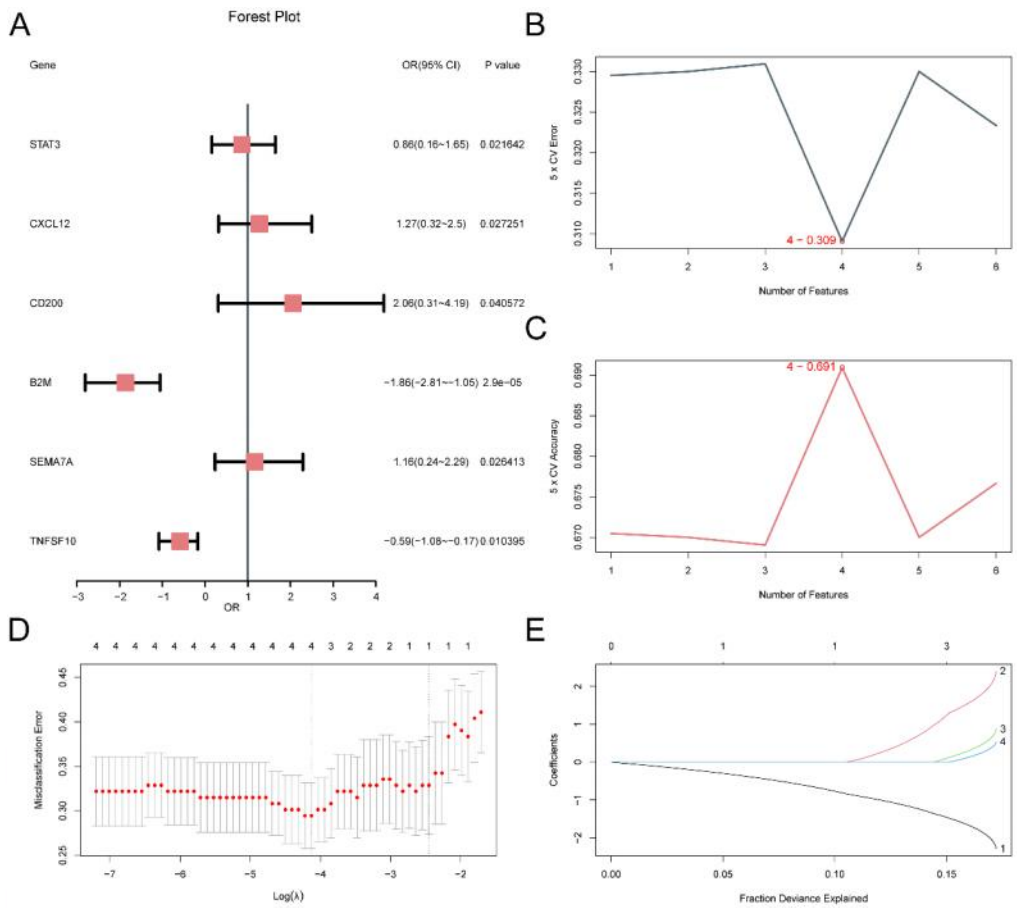
参考文献

[1] Bos LDJ, Ware LB. Acute respiratory distress syndrome: causes, pathophysiology, and phenotypes[J]. Lancet, 2022, 400(10358): 1145-1156.

[2] Bellani G, Laffey JG, Pham T, et al. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries[J]. JAMA, 2016, 315(8): 788-800.
[3] Huppert LA, Matthay MA, Ware LB. Pathogenesis of acute respiratory distress syndrome[J]. Semin Respir Crit Care Med, 2019, 40(1): 31-39.
[4] Aegerter H, Lambrecht BN, Jakubzick CV. Biology of lung macrophages in health and disease[J]. Immunity, 2022, 55(9): 1564-1580.
[5] Yu WJ, Yang ML, Lv BW, et al. CD40L-activated DC promotes Th17 differentiation and inhibits Th2 differentiation in sepsis-induced lung injury via cGAS-STING signaling[J]. Biochem Genet, 2025, 63(3): 2455-2469.
[6] Zhong XH, Jin JY, Zhang H, et al. Group 2 innate lymphoid cells derived IL-9 reduces macrophage apoptosis and attenuates acute lung injury in sepsis[J]. Inflamm Res, 2025, 74(1): 38.
[7] Cheng L, Jiao Y, Jiang W, et al. IL-33 deficiency attenuates lung inflammation by inducing Th17 response and impacting the Th17/treg balance in LPS-induced ARDS mice via dendritic cells[J]. J Immunol Res, 2022, 2022: 9543083.
[8] Guervilly C, Fournier T, Chommeloux J, et al. Ultra-lung-protective ventilation and biotrauma in severe ARDS patients on veno-venous extracorporeal membrane oxygenation: a randomized controlled study[J]. Crit Care, 2022, 26(1): 383.
[9] Sun MF, Yang QQ, Hu CL, et al. Identification and validation of autophagy-related genes in sepsis-induced acute respiratory distress syndrome and immune infiltration[J]. J Inflamm Res, 2022, 15: 2199-2212.
[10] Wu XL, Guo YN. Role of cellular senescence genes and immune infiltration in sepsis and sepsis-induced ARDS based on bioinformatics analysis[J]. J Inflamm Res, 2024, 17: 9119-9133.
[11] Liu TT, Gao L, Li XY. Identification of diagnostic biomarkers and therapeutic targets in sepsis-associated ARDS via combining bioinformatics with machine learning analysis [J]. J Inflamm Res, 2025, 18: 9523-9536.
[12] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update [J]. Nucleic Acids Res, 2013, 41(Database issue): D991-D995.
[13] Gorman EA, O'Kane CM, McAuley DF. Acute respiratory distress syndrome in adults: diagnosis, outcomes, long-term sequelae, and management[J]. Lancet, 2022, 400(10358): 1157-1170.
[14] Yang XT, Wang J, Liu W. Molecular markers of type Ⅱ alveolar epithelial cells in acute lung injury by bioinformatics analysis[J]. Sci Rep, 2023, 13(1): 17797.
[15] Liang QC, Zhou Q, Li JH, et al. Validation of novel hub genes and molecular mechanisms in acute lung injury using an integrative bioinformatics approach [J]. Acta Biochim Biophys Sin, 2021, 53(3): 342-353.
[16] Ming TQ, Dong MY, Song XM, et al. Integrated analysis of gene co-expression network and prediction model indicates immune-related roles of the identified biomarkers in sepsis and sepsis-induced acute respiratory distress syndrome[J]. Front Immunol, 2022, 13: 897390.
[17] Wang KT, Yang JL, Deng JP, et al. Pinocembrin reduces pyroptosis to improve flap survival by modulating the TLR4/NF-κB/NLRP3 signaling pathway[J]. Biochim Biophys Acta Mol Basis Dis, 2025, 1871(3): 167710.
[18] Huang HR, Cho SJ, Harris RM, et al. RIPK3 activates MLKL-mediated necroptosis and inflammasome signaling during Streptococcus Infection[J]. Am J Respir Cell Mol Biol, 2021, 64(5): 579-591.
[19] Liu YL, Zhu H, Chen H, et al. LPS-induced TMBIM6 splicing drives endothelial necroptosis and aggravates ALI[J]. Respir Investig, 2025, 63(2): 191-199.
[20] Leavy OC, Kawano-Dourado L, Stewart ID, et al. Rheumatoid arthritis and idiopathic pulmonary fibrosis: a bidirectional Mendelian randomisation study[J]. Thorax, 2024, 79(6): 538-544.
[21] Zhao R, Zhang YW, Guo JC, et al. Genetic evidence reveals a causal relationship between rheumatoid arthritis and interstitial lung disease[J]. Front Genet, 2024, 15: 1395315.
[22] Peng JN, Tang R, He J, et al. S1PR3 inhibition protects against LPS-induced ARDS by inhibiting NF-κB and improving mitochondrial oxidative phosphorylation[J]. J Transl Med, 2024, 22(1): 535.
[23] Lee SE, Kim IH, Kang YC, et al. Mitochondrial transplantation attenuates lipopolysaccharide-induced acute respiratory distress syndrome[J]. BMC Pulm Med, 2024, 24(1): 477.
[24] Li DP, Wang Y, Tang L, et al. CD200-CD200R1 signalling attenuates imiquimod-induced psoriatic inflammation by inhibiting the activation of skin inflammatory macrophages[J]. Int Immunopharmacol, 2020, 78: 106046.
[25] Patoine D, Bouchard K, Lemay AM, et al. Specificity of CD200/CD200R pathway in LPS-induced lung inflammation[J]. Front Immunol, 2022, 13: 1092126.
[26] Li X, Xie WL, Pan Q, et al. Semaphorin 7A interacts with nuclear factor NF-kappa-B p105 via integrin β1 and mediates inflammation [J]. Cell Commun Signal, 2023, 21(1): 24.
[27] Granja T, Köhler D, Tang LY, et al. Semaphorin 7A coordinates neutrophil response during pulmonary inflammation and sepsis[J]. Blood Adv, 2024, 8(11): 2660-2674.
[28] Kubo M, Harada Y, Sasaki T. The role of dendritic cells in the instruction of helper T cells in the allergic March[J]. Int Immunol, 2024, 36(11): 559-566.
[29] Gulubova MV, Valkanov SP, Ignatova MMK, et al. Tertiary lymphoid structures in colorectal cancer - organization and immune cell interactions[J]. Am J Clin Exp Immunol, 2024, 13(6): 236-245.
[30] Phanthunane C, Wijers R, De Herdt MJ, et al. Intratumoral niches of B cells and follicular helper T cells, and the absence of regulatory T cells, associate with longer survival in early-stage oral tongue cancer patients[J]. Cancers, 2022, 14(17): 4298.
[31] Li C, Yu S, Chen J, et al. Risk stratification based on DNA damage-repair-related signature reflects the microenvironmental feature, metabolic status and therapeutic response of breast cancer [J]. Front Immunol, 2023, 14: 1127982.
[32] Liu GH, Wang LP, Ji LL, et al. Identifying prognostic markers in spatially heterogeneous breast cancer microenvironment [J]. J Transl Med, 2023, 21(1): 580.
[33] Cruz FM, Orellano LAA, Chan A, et al. Alternate MHC I antigen presentation pathways allow CD8$^+$ T-cell recognition and killing of cancer cells in the absence of β2M or TAP[J]. Cancer Immunol Res, 2025, 13(1): 98-108.
[34] Han XW, Zhang JY, Li WD, et al. The role of B2M in cancer immunotherapy resistance: function, resistance mechanism, and reversal strategies[J]. Front Immunol, 2025, 16: 1512509.
[35] Dar AA, Patil RS, Pradhan TN, et al. Myeloid-derived suppressor cells impede T cell functionality and promote Th17 differentiation in oral squamous cell carcinoma[J]. Cancer Immunol Immunother, 2020, 69(6): 1071-1086.
[36] Pan PY, Ma G, Weber KJ, et al. Immune stimulatory receptor CD40 is required for T-cell suppression and T regulatory cell activation mediated by myeloid-derived suppressor cells in cancer[J]. Cancer Res, 2010, 70(1): 99-108.
[37] Song C, Li HT, Mao Z, et al. Delayed neutrophil apoptosis may enhance NET formation in ARDS[J]. Respir Res, 2022, 23(1): 155.
[38] Huang JH, Zhao CM, Zhang SW. Semaphorin 7A promotes endothelial permeability and inflammation via plexin C1 and integrin β1 in Kawasaki disease[J]. BMC Pediatr, 2024, 24(1): 285.

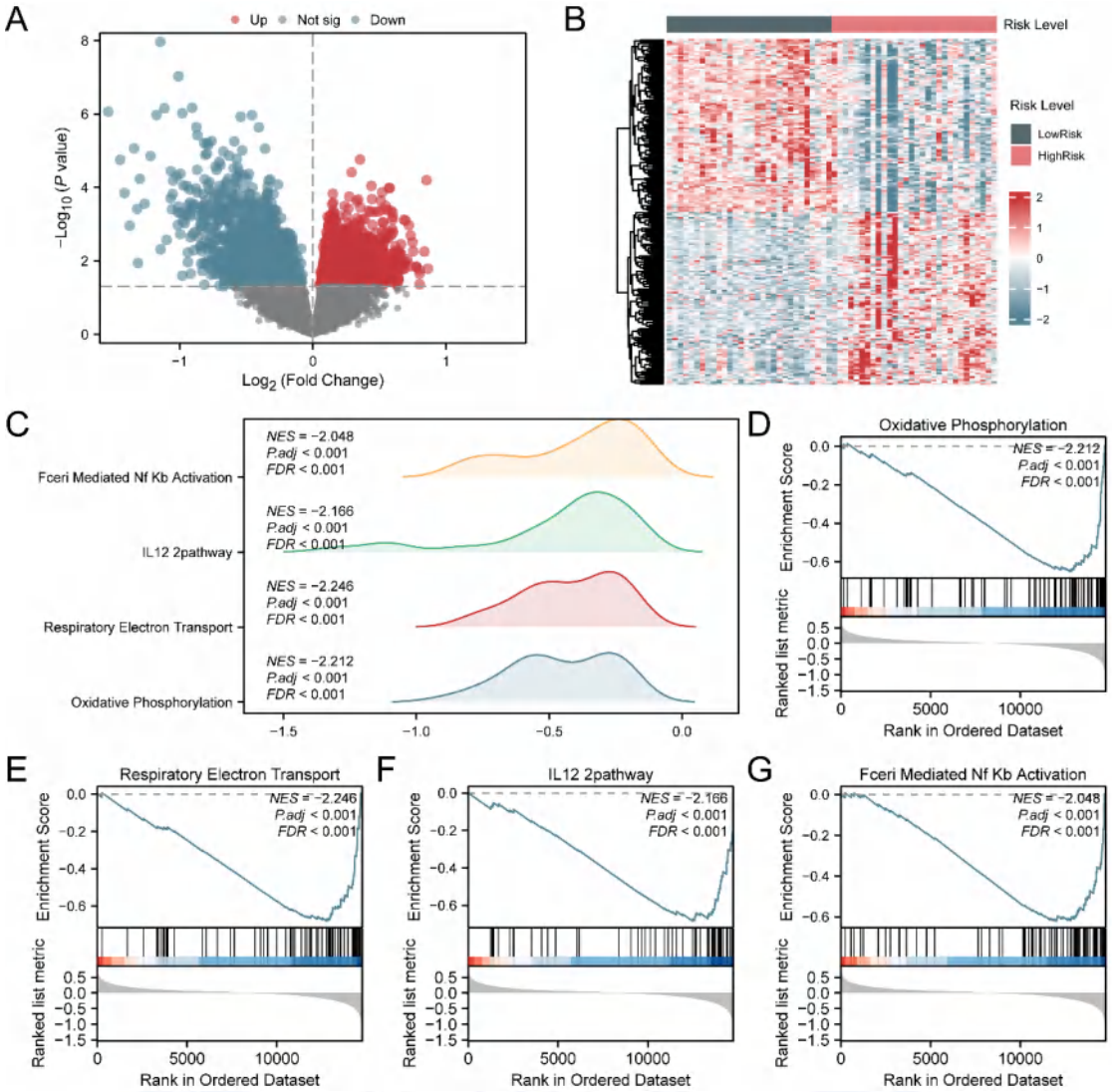
收稿日期:2025-12-23　**修回日期**:2026-03-10　**编辑**:石嘉莹

Supplementary figures:



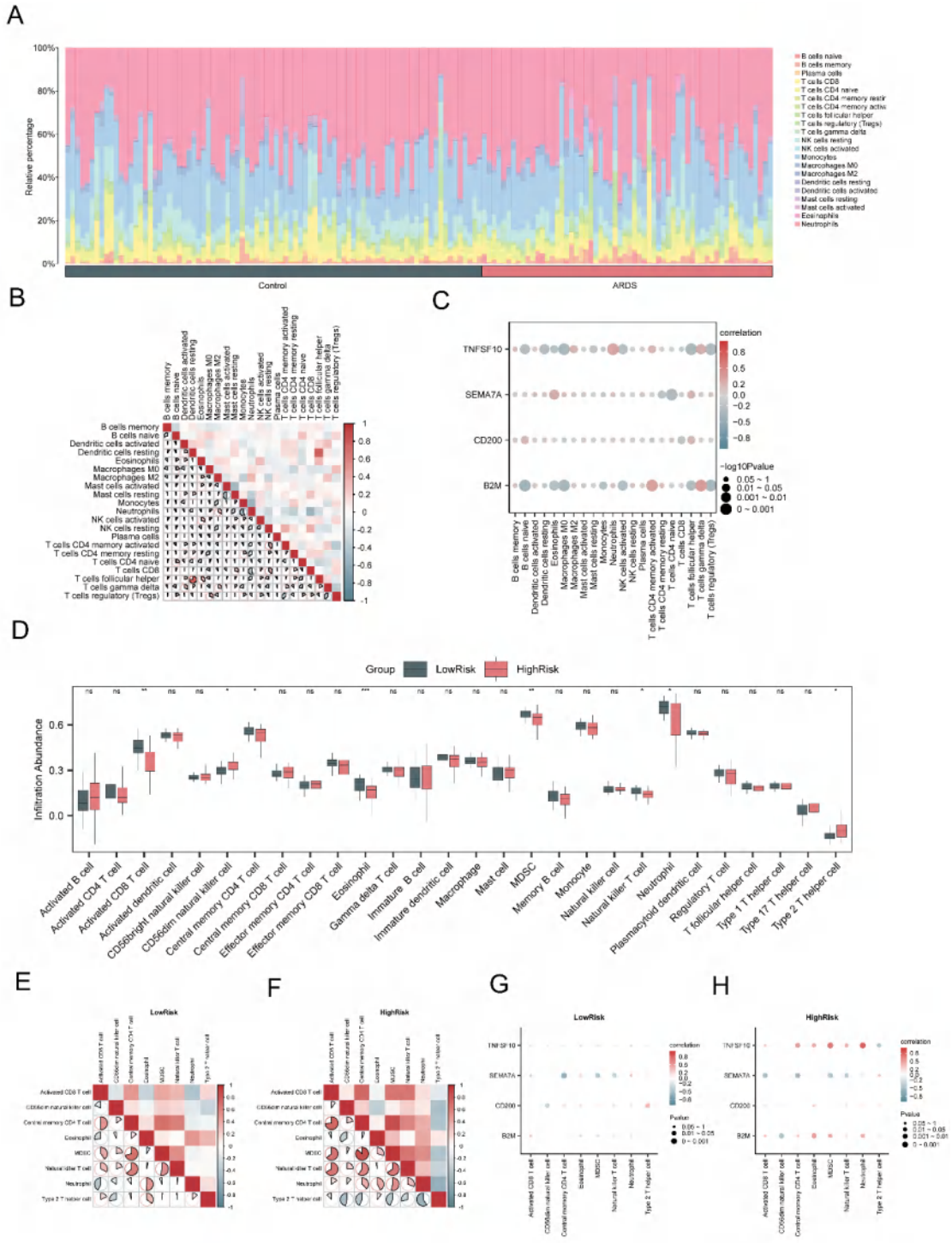
Note: A, Forest plot of module genes in logistic regression analysis; B-C,Optimal number of genes determined by the random forest algorithm based on error rate and accuracy; D-E, Diagnostic model plot and variable trajectory plot based on least absolute shrinkage and selection operator (LASSO) regression.

Supplementary Fig.1 Construction process of the ARDS diagnostic model



Note: A, volcano plot of differentially expressed genes; B, heatmap of differential gene expression; C, GSEA ridge plot of four biological functions; D-G, GSEA results showing significant enrichment of high ARDS risk in oxidative phosphorylation (D), respiratory electron transport (E), interleukin-12 (IL-12) signaling pathway (F), and Fc ε RI-mediated nuclear factor- κ B (NF- κ B) activation (G).

Supplementary Fig.2 Differential analysis and GSEA enrichment results between high- and low-risk groups



Note: A, immune cell proportion plot; B, correlation heatmap of immune cells; C, bubble plot showing the correlation between immune cell infiltration abundance and model genes; D, comparison of immune cells between the high- and low-risk groups; E-F, correlation heatmaps of immune cell infiltration abundance in the low-risk group (E) and high-risk group (F); G-H, bubble plots showing the correlation between immune cell infiltration abundance and model genes in the low-risk group (G) and high-risk group (H).

Supplementary Fig.3 Immune infiltration analysis results based on CIBERSORT and ssGSEA