



OPEN ACCESS

EDITED BY

Pilar Giraldo,
University of Zaragoza, Spain

REVIEWED BY

Jinmei Luo,
Third Affiliated Hospital of Sun Yat-sen
University, China
Pedro Martínez-Paz,
University of Valladolid, Spain

*CORRESPONDENCE

Shaodan Feng,
✉ sdfeng@fjmu.edu.cn
Xinlan Zhang,
✉ sy2024010@fjtcu.edu.cn

[†]These authors have contributed equally to this work and share first authorship

RECEIVED 15 January 2025

ACCEPTED 10 July 2025

PUBLISHED 22 July 2025

CITATION

Lin C, Zheng M, Wu W, Wang Z, Lu G, Feng S and Zhang X (2025) Development and validation of a machine-learning-based model for identification of genes associated with sepsis-associated acute kidney injury.
Front. Genet. 16:1561331.
doi: 10.3389/fgene.2025.1561331

COPYRIGHT

© 2025 Lin, Zheng, Wu, Wang, Lu, Feng and Zhang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Development and validation of a machine-learning-based model for identification of genes associated with sepsis-associated acute kidney injury

Chen Lin ^{1†}, Meng Zheng ^{2†}, Wensi Wu ¹,
Zhishan Wang ¹, Guofeng Lu ¹, Shaodan Feng ^{3*} and
Xinlan Zhang ^{1*}

¹Department of Emergency, The Third Affiliated People's Hospital, Fujian University of Traditional Chinese Medicine, Fuzhou, China, ²Hemodialysis Center, The Third Affiliated People's Hospital, Fujian University of Traditional Chinese Medicine, Fuzhou, China, ³Department of Emergency, The First Affiliated Hospital, Fujian Medical University, Fuzhou, China

Background: Sepsis frequently induces acute kidney injury (AKI), and the complex interplay between these two conditions worsens prognosis, prolongs hospitalization, and increases mortality. Despite therapeutic options such as antibiotics and supportive care, early diagnosis and treatment remain a challenge. Understanding the underlying molecular mechanisms linking sepsis and AKI is critical for the development of effective diagnostic tools and therapeutic strategies.

Methods: We used two sepsis (GSE57065 and GSE28750) and three AKI (GSE30718, GSE139061, and GSE67401) datasets from the NCBI Gene Expression Omnibus (GEO) for model development and validation, and performed batch effect mitigation, differential gene, and functional enrichment analysis using R software packages. We assessed 113 combinations of 12 different algorithms to develop an internally and externally validated machine-learning model for diagnosing AKI. Finally, we used functional enrichment analysis to identify potential therapeutic agents for AKI.

Results: We identified 556 and 725 DEGs associated with sepsis and AKI, respectively, with 28 overlapping genes suggesting shared pathways. Functional enrichment analysis revealed important associations of AKI with immune responses and cell adhesion processes. The immune infiltration analysis showed significant differences in immune cell presence between sepsis and AKI patients compared with the control group. The machine-learning models identified eight key genes (*NR3C2*, *PLEKHO1*, *CEACAM1*, *CDC25B*, *HEPACAM2*, *VNN1*, *SLC2A3*, *RPL36*) with potential for diagnosing AKI. The diagnostic performance of the model constructed in this way was excellent (area under the curve = 0.978), especially in the under 60 years and male patient subgroups. The diagnostic performance outperformed previous models in both the training and validation sets. In addition, cyclosporin A and nine other drugs were identified as potential agents for treating sepsis-associated AKI.

Conclusion: This study highlights the potential of integrating bioinformatics and machine-learning approaches to generate a new diagnostic model for sepsis-associated AKI using molecular crossovers with sepsis. The genes identified have potential to serve as biomarkers and therapeutic targets, providing avenues for future research aimed at enhancing sepsis-associated AKI diagnosis and treatment.

KEYWORDS

sepsis, acute kidney injury, machine learning, diagnostic modeling, immune infiltration

1 Introduction

Sepsis and acute kidney injury (AKI) are serious healthcare challenges that place a heavy burden on patient management and healthcare systems worldwide (Zarjou and Agarwal, 2011; Skube et al., 2018). Every year, millions of people develop sepsis, leading to increased mortality rates and healthcare costs as a result of longer hospital stays and increased need for intensive care (Souza et al., 2021). Current treatment strategies for sepsis-associated AKI (SA-AKI) include antibiotic therapy, fluid resuscitation, and supportive care (Kellum et al., 2019; Montomoli et al., 2019; Reddy et al., 2023). However, these approaches are often limited by the difficulty of early detection and accurate prediction of patient prognosis. The complex biological interactions between sepsis and AKI highlight the urgency of further research to unravel the underlying mechanisms and identify novel biomarkers that can aid in early diagnosis and timely intervention. The purpose of this study was to explore the relationship between sepsis and AKI in order to obtain new insights to improve the early diagnosis and treatment of patients with SA-AKI.

Previous studies indicate a biological link between sepsis and AKI, suggesting sepsis may trigger AKI development (Martensson et al., 2010; Liu et al., 2024). Conversely, AKI can exacerbate the clinical severity of sepsis (Vodovar et al., 2021). Determining the time-point of onset of AKI in sepsis is difficult (Murugan et al., 2010; Wen et al., 2020). To enhance early diagnosis and treatment of SA-AKI, a deeper understanding of the genes and mechanisms linking these conditions is vital, and biomarkers play an important role in this process (Derive and Gibot, 2011; Odum et al., 2022). Traditional kidney function markers like serum creatinine and urine output are essential for diagnosing AKI. Their limitations are becoming apparent, especially in patients with sepsis, where changes in kidney function tend to lag behind the AKI. This phenomenon has prompted clinicians to search for new biomarkers in order to recognize AKI early and improve the prognosis. Binding of tissue inhibitory factor-2 (TIMP-2) and insulin-like growth factor binding protein 7 (IGFBP7) in urine has been identified as a useful biomarkers for predicting SA-AKI (Su et al., 2017; Jiang and Zheng, 2022). In addition, the use of novel biomarkers such as kidney injury molecule-1 (KIM-1) and neutrophil gelatinase-associated lipid carrier (NGAL) provides more options for early diagnosis (Ersavas et al., 2016; Ucakturk et al., 2016; Laurentino et al., 2019). However, the diagnostic potential and immunologic significance of predicting AKI using sepsis genes remains largely unexplored.

Therefore, we chose to use bioinformatics methods and multiple machine learning strategies to identify key biomarkers associated

with SA-AKI, to elucidate the intrinsic mechanisms and linkages between these key biomarkers, to establish an early diagnostic model, and to provide new perspectives for the early identification of SA-AKI.

2 Methods

2.1 Data collection and consolidation

In this study, we obtained relevant datasets from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) using the keywords “sepsis” and “acute kidney injury.” We used two sepsis datasets (GSE57065, GSE28750) and three AKI datasets (GSE30718, GSE139061, GSE67401). GSE57065 contained data on 25 control and 82 sepsis samples; GSE28750 contained data on 20 control and 21 sepsis samples; GSE30718 contained data on 8 control and 28 AKI samples; and GSE139061 contained data on 9 control and 39 AKI samples. GSE67401 contained data on 58 control and 53 AKI samples (Table 1). Two sepsis datasets (GSE57065 and GSE28750) and two AKI datasets (GSE30718 and GSE139061) were merged as training sets. The “ComBat” function within the “sva package” of the R programming language was used to eradicate the batch effect from the data (Leek et al., 2012). The Institutional Review Board of the Third People’s Hospital, which is associated with Fujian University of Traditional Chinese Medicine, exempted the necessity for ethical approval. This exemption was granted due to the fact that the data utilized is readily accessible to the public via the GEO database. Consequently, participants, along with their legal guardians or next of kin, were not mandated to furnish written informed consent for their involvement in this study, in accordance with relevant national regulations and institutional protocols.

2.2 Identification of differentially expressed genes (DEGs)

We used the R package “limma” to identify differentially expressed genes (DEGs) in the sepsis and AKI datasets (Ritchie et al., 2015). The established thresholds were $|\log_2 \text{ fold change (FC)}| > 1$ and $P < 0.05$. To illustrate the findings, volcano plots and heat maps were created using the “ggplot2” and “pheatmap” packages in R. We constructed a Venn diagram to show the overlap between the DEGs associated with sepsis and those associated with AKI.

TABLE 1 Basic information on the GEO dataset.

Dataset	Disease	Sample	Platform	Attribute
GSE57065	Sepsis	82 patients with sepsis and 25 controls	GPL570	Training set
GSE28750	Sepsis	21 patients with sepsis and 20 control	GPL570	Training set
GSE30718	AKI	28 patients with AKI and 8 controls	GPL570	Training set
GSE139061	AKI	39 patients with AKI and 9 controls	GPL20301	Training set
GSE67401	AKI	53 patients with AKI and 58 controls	GPL9115	Validation set

AKI: acute kidney injury.

2.3 Enrichment analysis of intersecting genes

We constructed a protein-protein interaction (PPI) network using the GeneMANIA (<http://genemania.org/>) website to show the interrelationships between crossover genes. For genomic functional enrichment analysis, we used the Gene Ontology (GO) and Kyoto Encyclopedia of the Genome (KEGG) with the R “org.Hs.e.g.,db” and “clusterProfiler” packages ($P < 0.05$).

2.4 Machine-learning algorithms

We evaluated the diagnostic accuracy of 113 combinations of 12 distinct machine-learning algorithms (Supplementary Table S1). We used the AKI de-batch effect datasets (GSE30718 and GSE139061) as the training set and an external dataset (GSE67401) as the validation set. During the model training phase, we evaluated the performance of a variety of machine-learning algorithms, including Enet regression ($\lambda = 0.1$), Lasso regression ($\lambda = 0.05$), Ridge regression ($\lambda = 1.0$), SVM ($C = 1.0$, $\gamma = 0.01$), LDA, Gradient Boosting Machine (GBM, learning rate = 0.1, number of trees = 100), RF (number of trees = 200), XGBoost (XGB, learning rate = 0.01, number of trees = 150). The models were developed using the training set and the hyperparameters were refined by cross-validation (Supplementary Table S2). During the model evaluation phase, the area under the curve (AUC) of each model was computed using the validation set to evaluate its classification accuracy. The model with the highest AUC was chosen as the optimal model. Calibration curves were used to evaluate the reliability, whereas decision curve analysis (DCA) was used to assess the clinical applicability of each model. The performance of the models was assessed in different subgroups and compared with published models of key genes for sepsis AKI pathogenesis were compared (DeLong test, $P < 0.05$).

2.5 Analysis of immune infiltration

We performed immune infiltration analysis of sepsis and AKI data using the “CIBERSORT” package in R with LM22 as the reference data set ($P < 0.05$). The results were illustrated using box-plots. Furthermore, heatmaps were used to show the correlation between immune cells and DEGs.

2.6 Identification of potential drug candidates

We used the Enrich platform (<https://maayanlab.cloud/Enrichr/>) to identify drug candidates capable of targeting the underlying pathological mechanisms associated with sepsis and AKI.

2.7 Statistical analysis

Data were processed and analyzed using R software (version 4.2.2; The R Foundation for Statistical Computing, Vienna, Austria). Normality of continuous variables was assessed using the Shapiro-Wilk test, supplemented by visual inspection of histograms. Variables conforming to a normal distribution (Shapiro-Wilk $P \geq 0.05$) were compared between groups using the independent samples t-test, whereas non-normally distributed variables (Shapiro-Wilk $P < 0.05$) were compared between groups using the Mann-Whitney U test. Statistical significance was defined as a two-tailed P value < 0.05 .

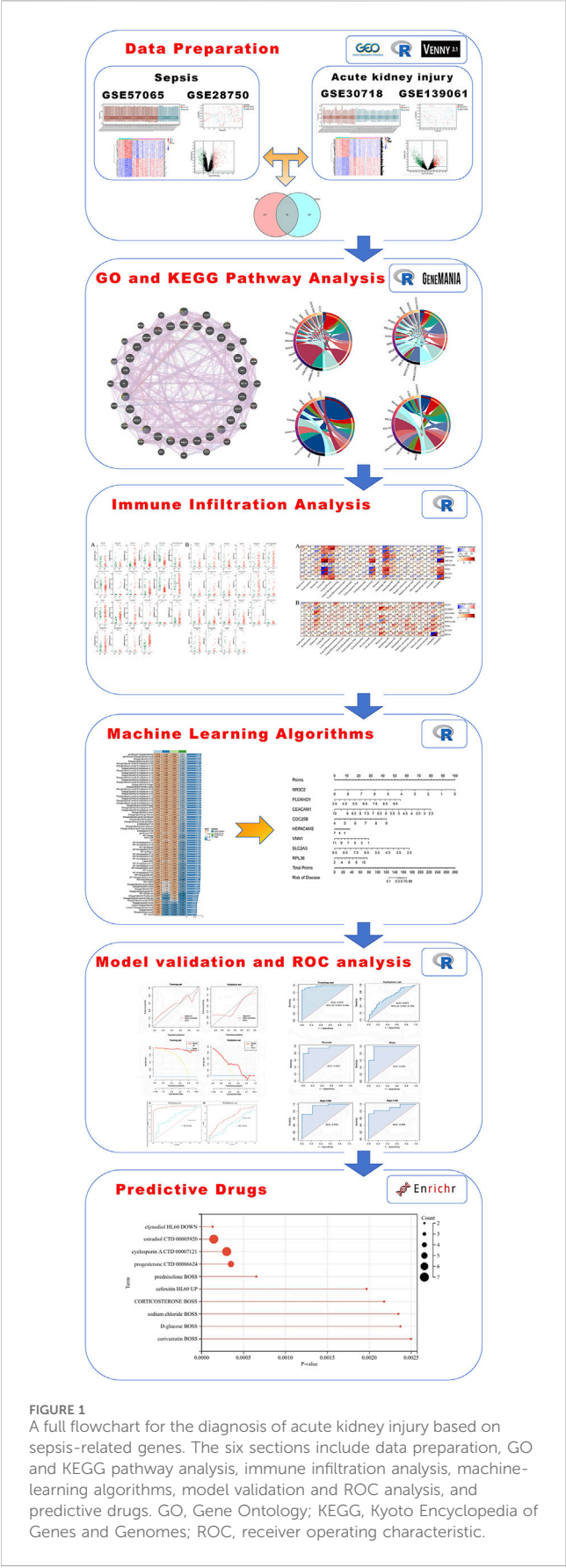
3 Results

3.1 Data collection and de-batching

A summary of the study design is shown in Figure 1. After eliminating the batch effects from the sepsis datasets (GSE57065 and GSE28750), 103 cases of sepsis and 45 healthy controls were identified (Figures 2A,B). Similarly, after eliminating the batch effects from the AKI datasets (GSE30718 and GSE139061), 67 cases of AKI and 17 healthy controls were identified (Figures 2C,D). Boxplots effectively highlighted the variances observed prior to and subsequent to the removal of the batch effect. Furthermore, Uniform Manifold Approximation and Projection (UMAP) plots indicated that, prior to the batch-effect correction, the samples from each dataset exhibited a central clustering pattern. However, after correction, the samples displayed a more dispersed distribution, indicating a successful mitigation of the batch effect (Figures 3A–D).

3.2 Determination of sepsis and AKI DEGs separately

A total of 309 upregulated DEGs and 247 downregulated DEGs were identified in the sepsis dataset (Figure 4A), and a total of



231 upregulated DEGs and 494 downregulated DEGs were identified in the AKI dataset (Figure 4B). Hierarchical clustering heatmaps of these DEGs revealed different expression patterns between the two groups (Figures 4C,D). Comparison of the two datasets revealed 28 overlapping co-morbid genes in sepsis and AKI (Figure 4E; Supplementary Table S3).

3.3 Functional enrichment analysis of sepsis and AKI intersection genes

A PPI network diagram depicting the genes implicated in SA-AKI was constructed using the GeneMANIA database (Figure 5A). The biological processes (BP) identified through GO enrichment analysis (Supplementary Table S4) encompassed the regulation of cell-cell adhesion facilitated by integrins, leukocyte activation, cell-cell adhesion mediated by integrins, innate immune responses, and myeloid leukocyte activation (Figure 5B). The cellular components (CC) comprised the tertiary granule membrane, secretory granule membrane, polysomal ribosome, secretory granule, and the bounding membranes of organelles (Figure 5C). The molecular functions (MF) included protein homodimerization activity, alpha-1,4-glucosidase activity, prostaglandin E receptor activity, pantetheine hydrolase activity, and receptor signaling protein tyrosine kinase activator activity (Figure 5D). Additionally, the KEGG enrichment analysis revealed pathways associated with prion diseases, the p53 signaling pathway, complement and coagulation cascades, Chaga's disease (American trypanosomiasis), and the cell cycle (Figure 5E; Supplementary Table S5).

3.4 Development of a diagnostic model for SA-AKI through machine learning

Model training was conducted by amalgamating data from the AKI datasets to establish a training set, using 113 combinations of 12 distinct machine-learning algorithms for performance evaluation. The final synthesis of the glmBoost and StepAim [both] algorithms yielded a model that demonstrated the best performance, with an AUC of 0.978. Validation against an external AKI dataset (GSE67401) indicated moderate performance with an AUC of 0.671 (Figures 6A–C; Supplementary Table S6). The model performed equally well for the subgroup under 60 years of age (AUC = 0.925) and the male subgroup (AUC = 0.956; Figures 6D–G). The algorithm identified eight key genes (*NR3C2*, *PLEKHO1*, *CEACAM1*, *CDC25B*, *HEPACAM2*, *VNN1*, *SLC2A3*, *RPL36*), which were subsequently used to develop a nomogram. The calibration and DCA curves demonstrated good agreement between the predicted and actual probabilities of occurrence, and good clinical utility of the model (Figures 7A–E).

3.5 Analysis of immune cell infiltration in sepsis and AKI

Immune infiltration analysis showed distinct infiltration levels in the sepsis group compared with the control group, which included

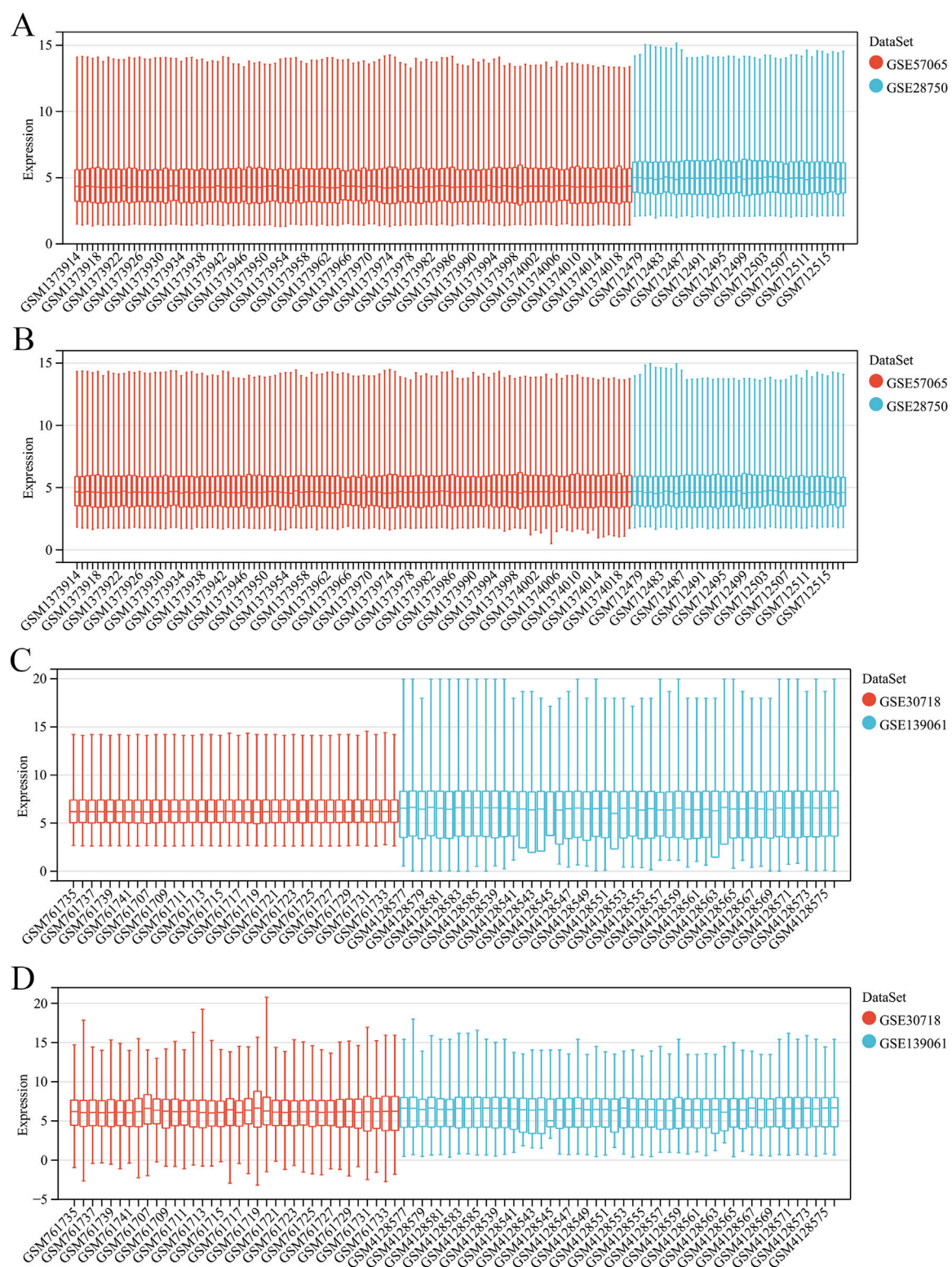
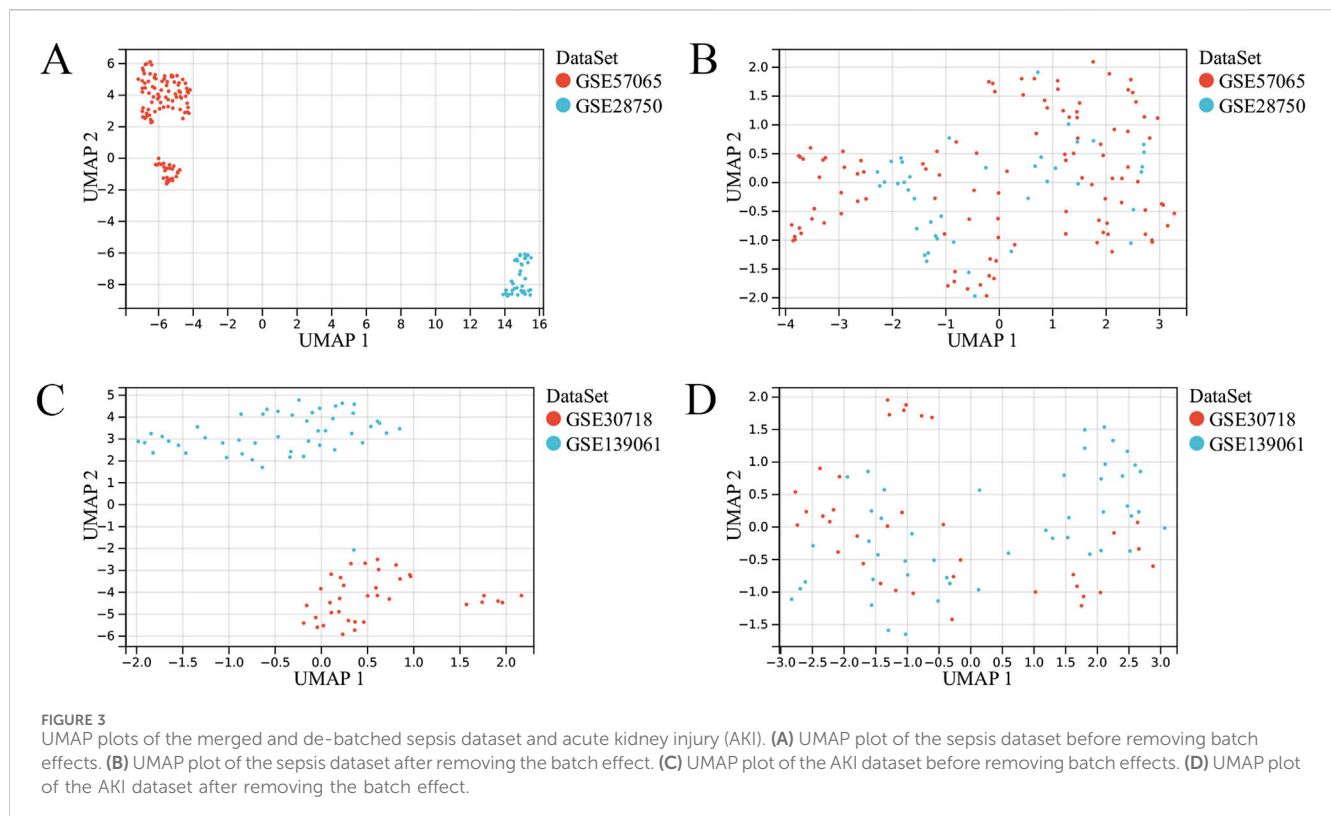


FIGURE 2

Merging and de-batching of sepsis dataset and acute kidney injury (AKI). (A) Boxplots of the sepsis dataset before removing batch effects. (B) Boxplots of the sepsis dataset after removing the batch effect. (C) Boxplots of the AKI dataset before removing batch effects. (D) Boxplots of the AKI dataset after removing the batch effect.



increased plasma cells, activated natural killer (NK) cells, M0 macrophages, eosinophils, CD8 T-cells, M1 macrophages, M2 macrophages, naive B cells, resting memory CD4 T-cells, naive CD4 T-cells, activated memory CD4 T-cells, gamma delta T cells, activated dendritic cells, neutrophils, and memory B cells in the sepsis group (Figure 8A). Similarly, differences in immune cell infiltration were observed between the AKI and control groups, including differences in plasma cells, CD8 T-cells, activated memory CD4 T-cells, monocytes, M0 macrophages, M1 macrophages, and eosinophils (Figure 8B).

Figure 9 shows the correlation between the immune cell counts and the sepsis and AKI model genes. Table 2 shows immune cell indicators were highly correlated (absolute magnitude of the correlation coefficient >0.60) with model genes in each data set.

3.6 Comparison of AKI diagnostic models

In order to comprehensively compare the performance of our model with other models, we compared the involvement of *VMP1*, *SLPI*, *PTX3*, *TIMP1*, *OLFM4*, *LCN2*, and *S100A9* genes in the development of AKI in patients with sepsis as described in the paper by Tang et al. (2021). The *VMP1* gene was not found in the validation cohort and was excluded. In the training cohort, our model achieved an AUC of 0.978 (95% CI: 0.952–0.995) vs 0.683 (95% CI: 0.538–0.828) for the Tang model (DeLong test, $P = 3.6 \times 10^{-5}$). The AUC of our model was also higher than that of the Tang model using the validation cohort with AUCs of 0.671 (95% CI: 0.575–0.767) and 0.470 (95% CI: 0.362–0.583) for our

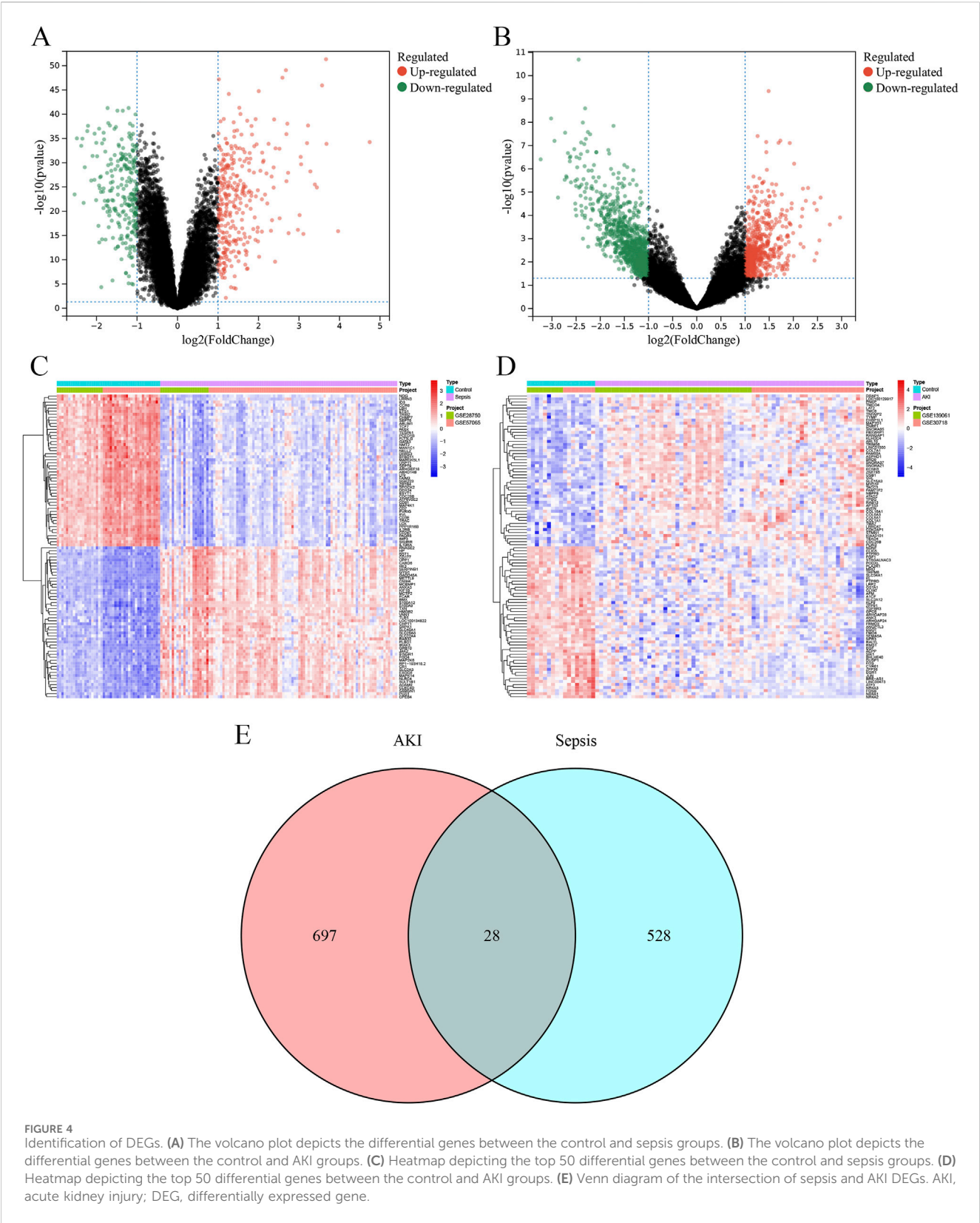
model and the Tang model, respectively (DeLong test, $P = 2.6 \times 10^{-4}$) (Figures 10A,B).

3.7 Drug predictions

We analyzed the model genes using the Drug Signatures Database (DSigDB) and the Enrich platform to identify potential pharmacological agents that could be used to treat AKI in patients with sepsis. The leading ten drug candidates identified were etynodiol (HL60 DOWN), estradiol (CTD 00005920), cyclosporin A (CTD 00007121), progesterone (CTD 00006624), prednisolone (BOSS), cefoxitin (HL60 UP), corticosteroid (BOSS), sodium chloride (BOSS), D-glucose (BOSS), and cerivastatin (BOSS) (Figure 11).

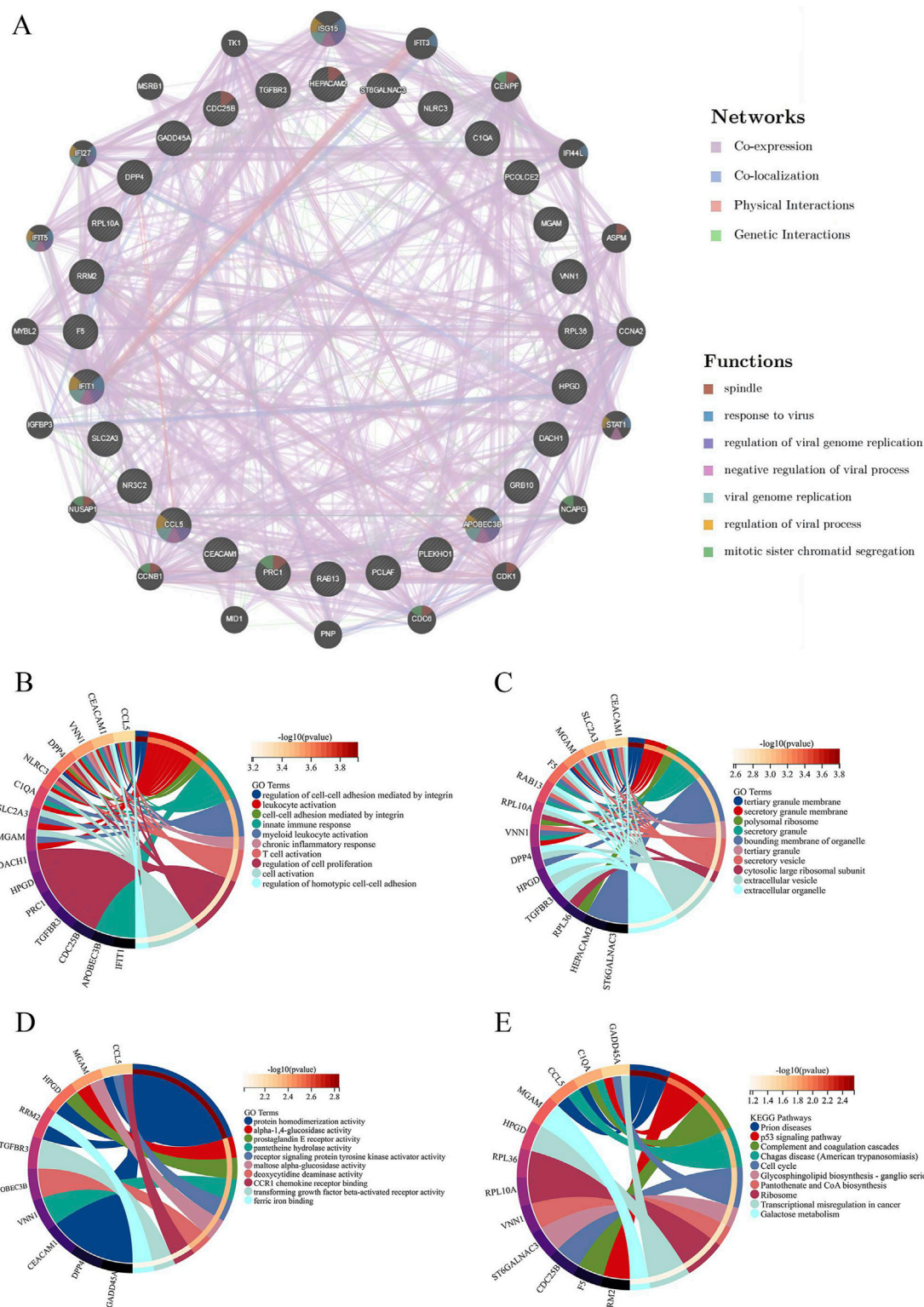
4 Discussion

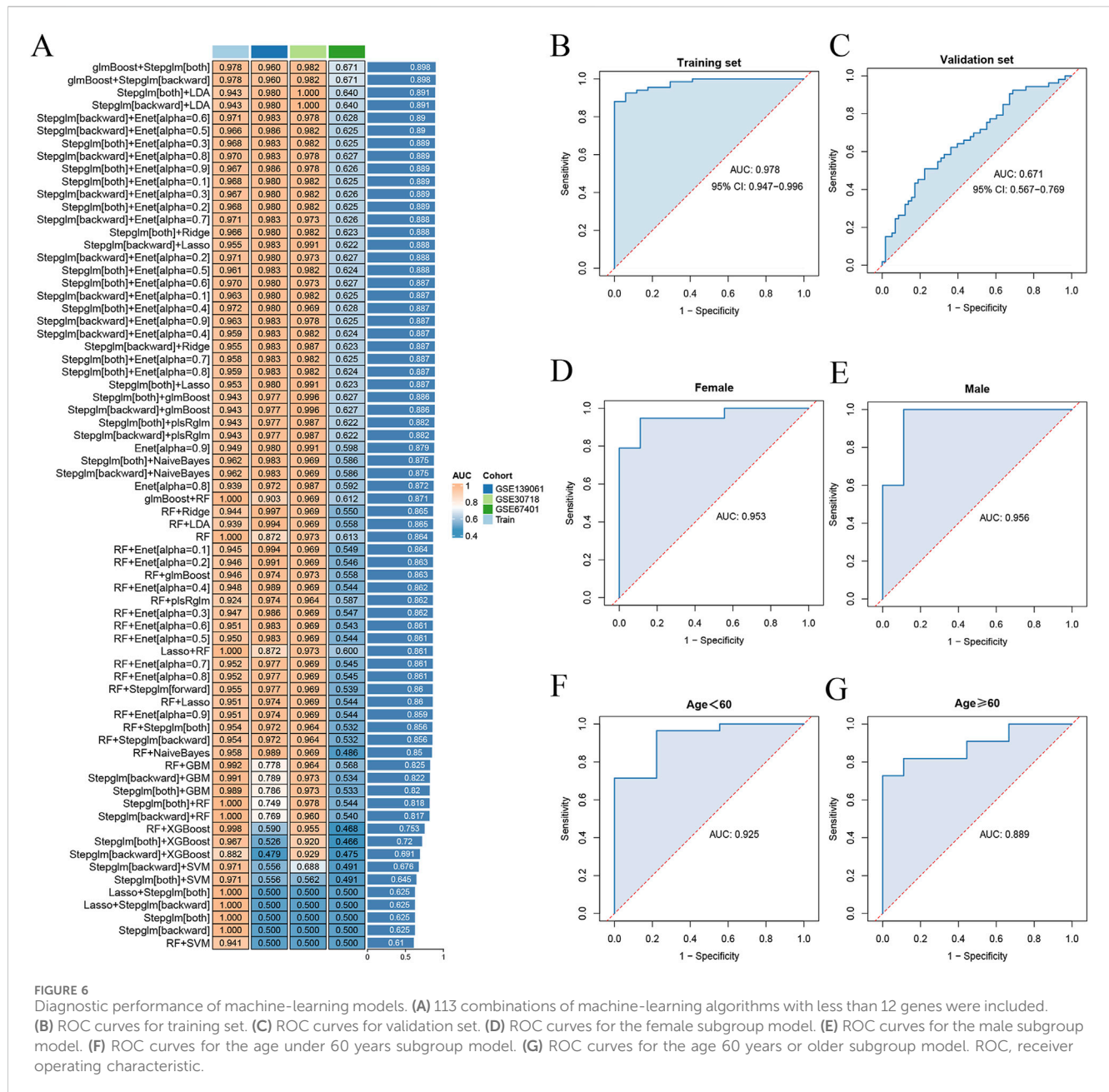
The integration of bioinformatics with machine learning is becoming progressively significant in the realm of disease diagnosis. With the rapid growth of biomedical data, traditional analytical methods are no longer adequate for a deeper understanding of complex biological systems (Qian et al., 2024). In recent years, many studies (Pillai et al., 2022; Alsaeedi et al., 2024) have shown that the combination of machine-learning methods with bioinformatics approaches can enhance the training and validation process, identify the best interpretable features, and facilitate deeper investigation of features and models. Therefore, this study developed



a model for diagnosing AKI based on a database of patients with sepsis, using a combination of machine learning and bioinformatics. The integration of these methods holds promise for improving the early detection and clinical treatment of SA-AKI.

This study highlights the intricate relationship between sepsis and AKI by identifying DEGs. 28 genes were identified as shared differential genes. These 28 overlapping genes may lead to common molecular pathways in the pathogenesis of both sepsis and AKI. This





is consistent with previous findings that emphasize the importance of changes in gene expression that emphasize the importance of changes in gene expression for understanding disease mechanisms (Goggi et al., 2020).

Through KEGG enrichment analysis, we identified several key pathways that reveal the intrinsic mechanisms of interdependence between sepsis and AKI. The p53 signaling pathway is crucial for the cellular response to stress, the induction of apoptosis, and the regulation of the cell cycle. Its dysregulation is associated with a variety of pathologic conditions, including sepsis (Hotchkiss et al., 2000). Activation of p53 may lead to kidney cell apoptosis and exacerbate AKI during sepsis (Lee et al., 2024; Lv et al., 2024). The complement and coagulation cascades represent an additional pathway of enrichment, aligning with findings from prior research (Lin et al., 2024; Wei et al., 2024). The complement system is activated during sepsis, promoting an inflammatory

response, and potentially leading to kidney microvascular injury. Activation of this pathway can lead to the formation of membrane attack complexes, which can damage kidney endothelial cells and further exacerbate AKI. The analysis revealed that the regulation of cell-cell adhesion mediated by integrins is an important biological process. Integrins are essential for maintaining the integrity of the endothelial barrier (Ayloo et al., 2022), and their dysregulation leads to increased vascular permeability and tissue edema (Infanger et al., 2008), a common feature of sepsis and AKI. Disruption of cell adhesion mechanisms may promote leukocyte infiltration and exacerbate inflammation, leading to AKI (Grande et al., 2010).

Immune cell analysis showed significant differences between the sepsis and control groups, highlighting the roles of specific immune cells in sepsis and AKI. The main cell types involved included activated NK cells, M0 macrophages, and CD8 T-cells. Activated NK

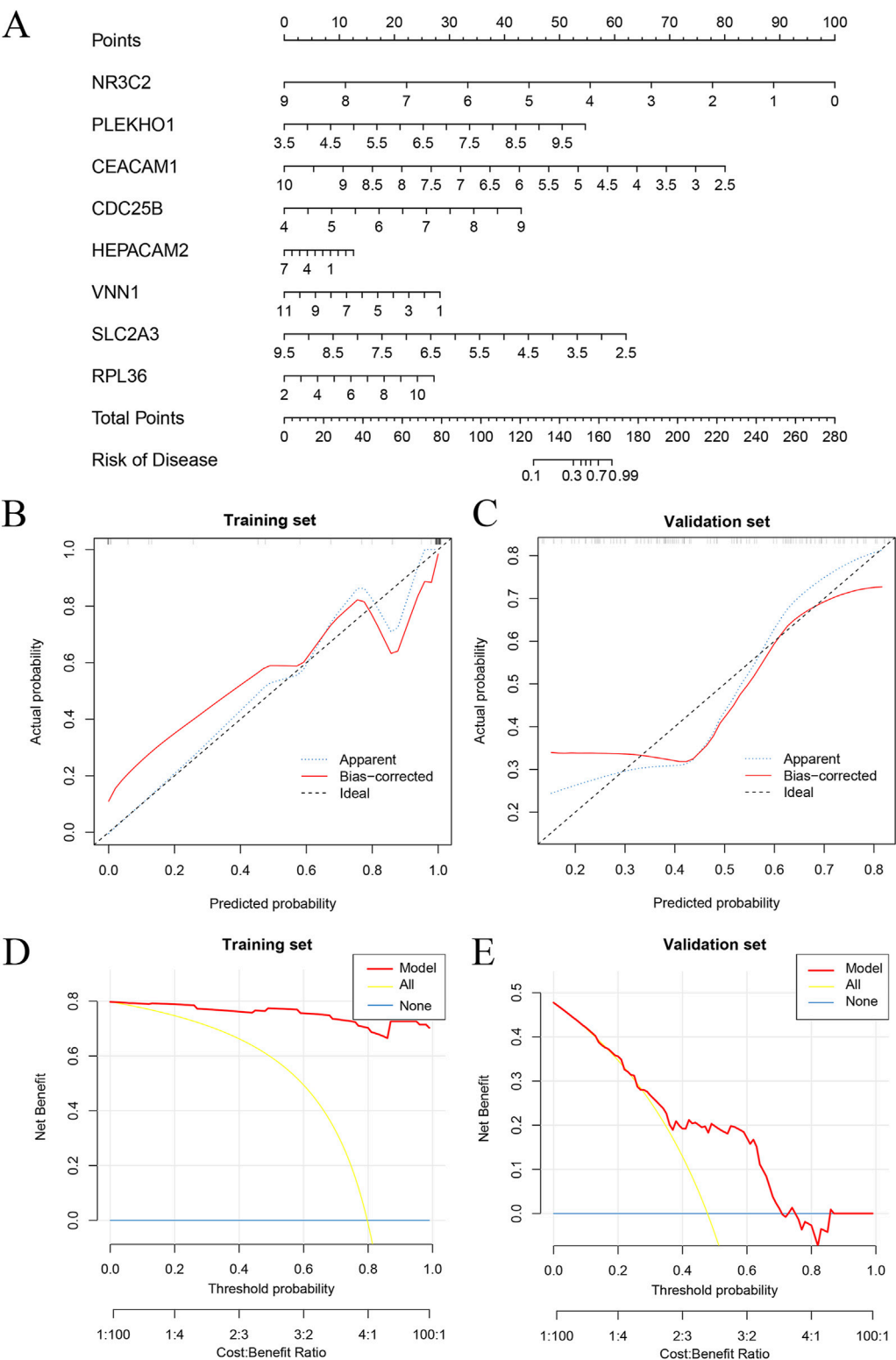
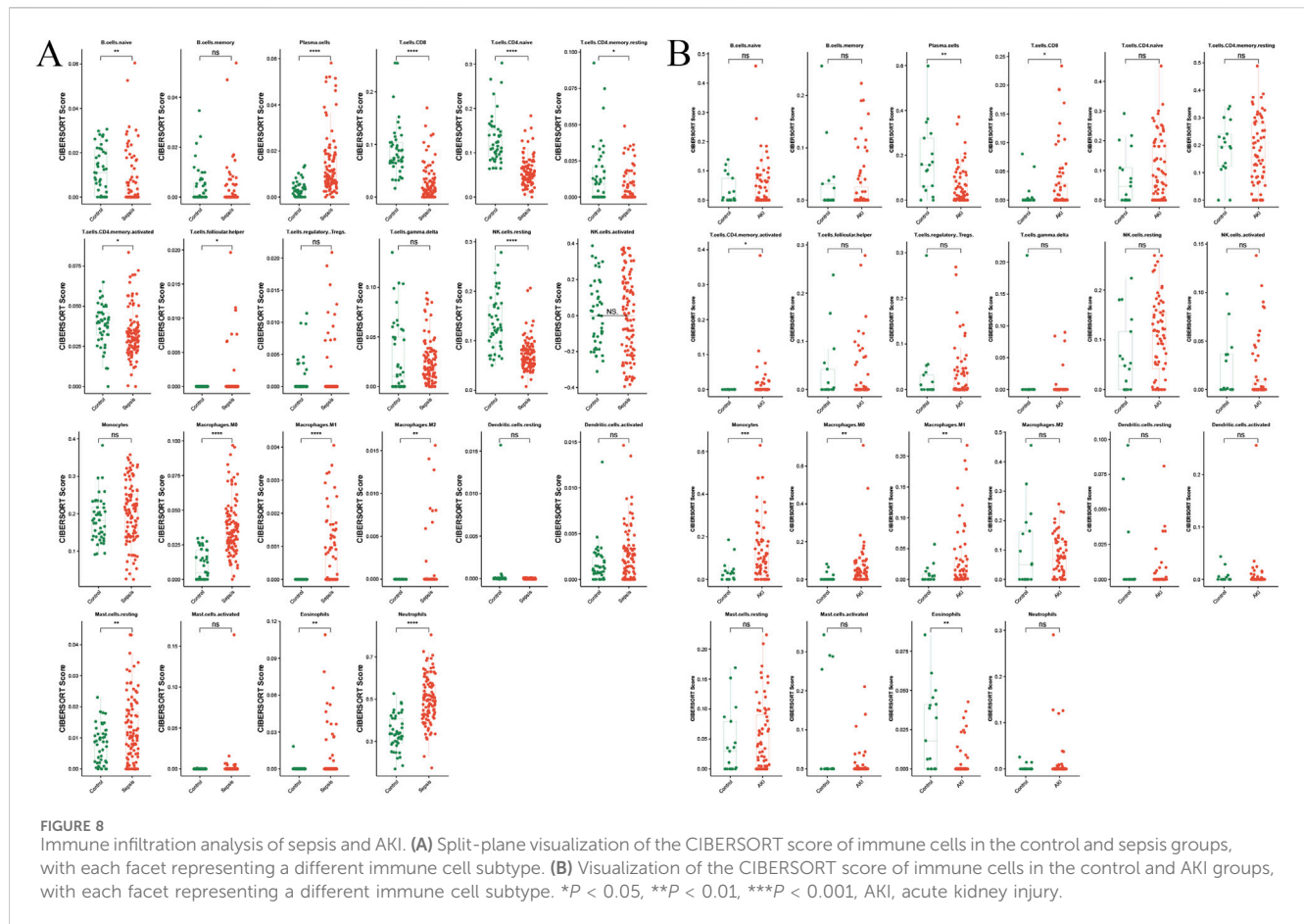


FIGURE 7 Nomogram and calibration of AKI diagnostic gene models. **(A)** Nomogram based on eight key genes. **(B)** Calibration curves for training set. **(C)** Calibration curves for validation set. **(D)** DCA curve for the training set. **(E)** DCA curve for the validation set. AKI, acute kidney injury; DCA, decision curve analysis.



cells are crucial for the innate immune response, recognizing and eliminating infected or malignant cells (Wu and Lanier, 2003; Brandstadter and Yang, 2011). In sepsis, the presence of NK cells is associated with an improved immune response to infection. NK cells produce pro-inflammatory cytokines, contributing to the systemic inflammatory response in sepsis (Jensen et al., 2021). The infiltration of activated NK cells in sepsis suggests that this may be a compensatory mechanism to restore homeostasis, but overactivation of NK cells can lead to tissue damage, including AKI. M0 macrophages maintain tissue homeostasis and initiate immune responses (Chen et al., 2017). The infiltration of M0 macrophages in sepsis suggests potential transformation into a pro-inflammatory phenotype, which is crucial in early sepsis. M0 macrophages can differentiate into inflammatory M1 or tissue-repairing M2 macrophages (Enderlin Vaz Da Silva et al., 2014; Isali et al., 2022). The balance between these phenotypes is critical in determining the outcome of sepsis and AKI. CD8 T-cells, which are cytotoxic, are key to the adaptive immune response (Jung et al., 2021). CD8 T-cells exacerbate inflammation and lead to organ dysfunction (Martin et al., 2018). In this study, the activation of CD8 T-cells was correlated with poorer prognosis of sepsis. The pattern of immune cell infiltration highlights the complex relationship between the immune system and sepsis and AKI. The study of activated NK cells, M0 macrophages, and CD8 T-cells could shed light on the pathogenesis of AKI in

patients with sepsis and provide further insights into the relationship between these two conditions.

Early diagnosis of SA-AKI is challenging for clinicians (Zhang et al., 2024). Delayed diagnosis not only leads to deterioration of the disease, but also sometimes leads to the transformation of SA-AKI into chronic kidney disease in some patients (Heung and Chawla, 2014), resulting in increased medical and social costs. Although the application of machine-learning techniques in this field shows good prospects, it still faces various challenges such as data quality, model interpretability, and clinical applicability. In a study based on machine-learning algorithms (Li et al., 2023; Li et al., 2024), it was shown that a model built using electronic medical record data enabled accurate prediction of SA-AKI in patients admitted to the intensive care unit. Our study differs from previous studies in that we focus on the overlap of genes between sepsis and AKI, which may be the key genes in the development of sepsis into AKI, and in this way build a more suitable model. Among 113 combinations of 12 machine-learning algorithms, the models were screened for number of genes < 12 . Of the models considered, the glmBoost and StepAim [both] algorithms combined produced the highest AUC value, which suggests their potential utility in a clinical setting. This algorithm also reduces the bias caused by incomplete algorithm selection and subjective human selection (Liu et al., 2022; Qin et al., 2023). In terms of model effect validation, our model showed excellent

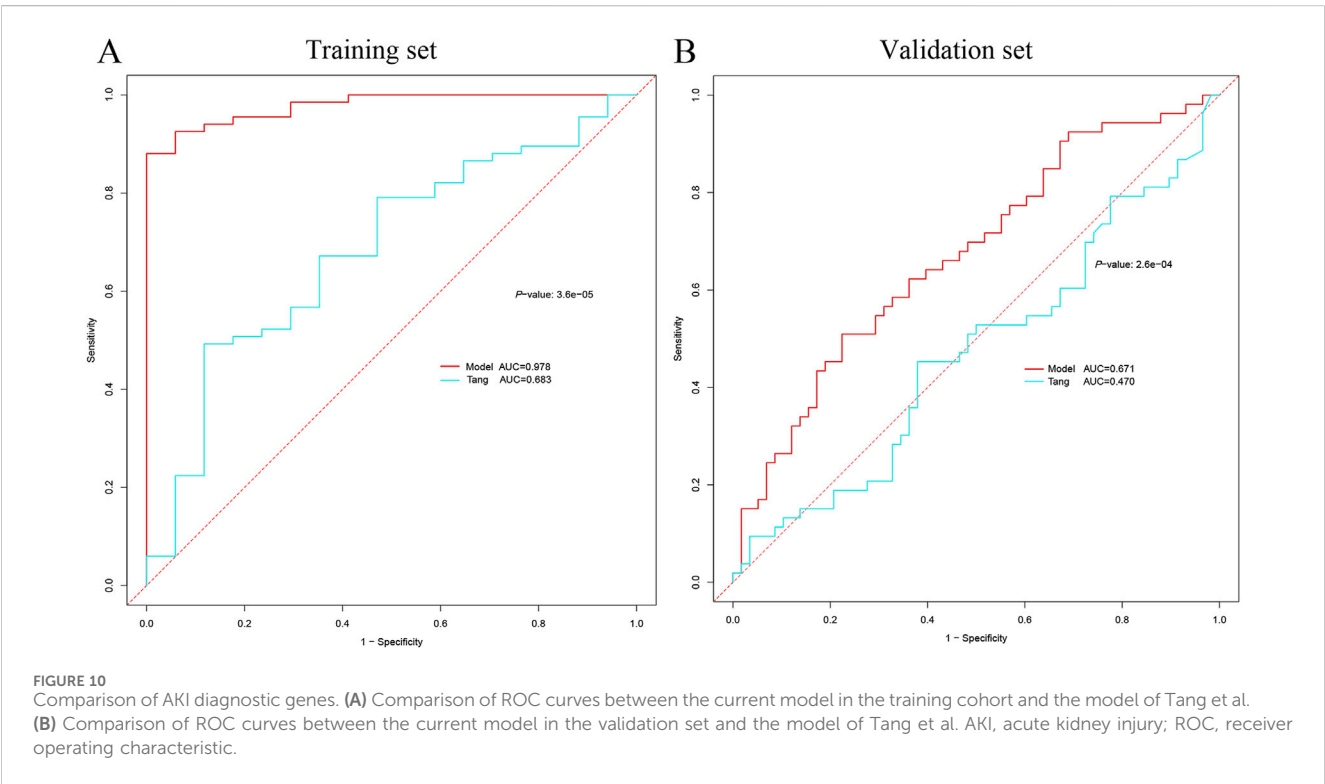
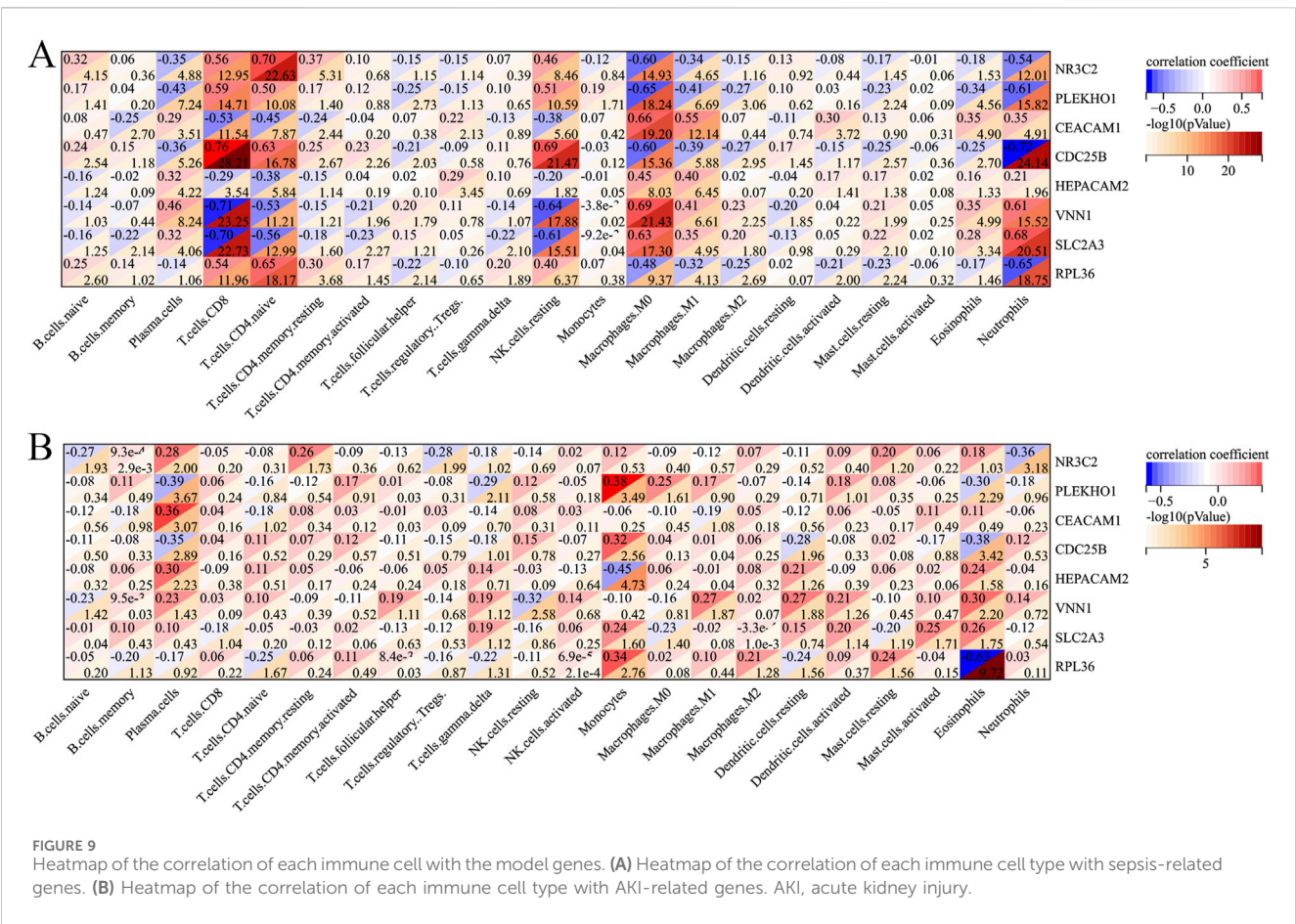
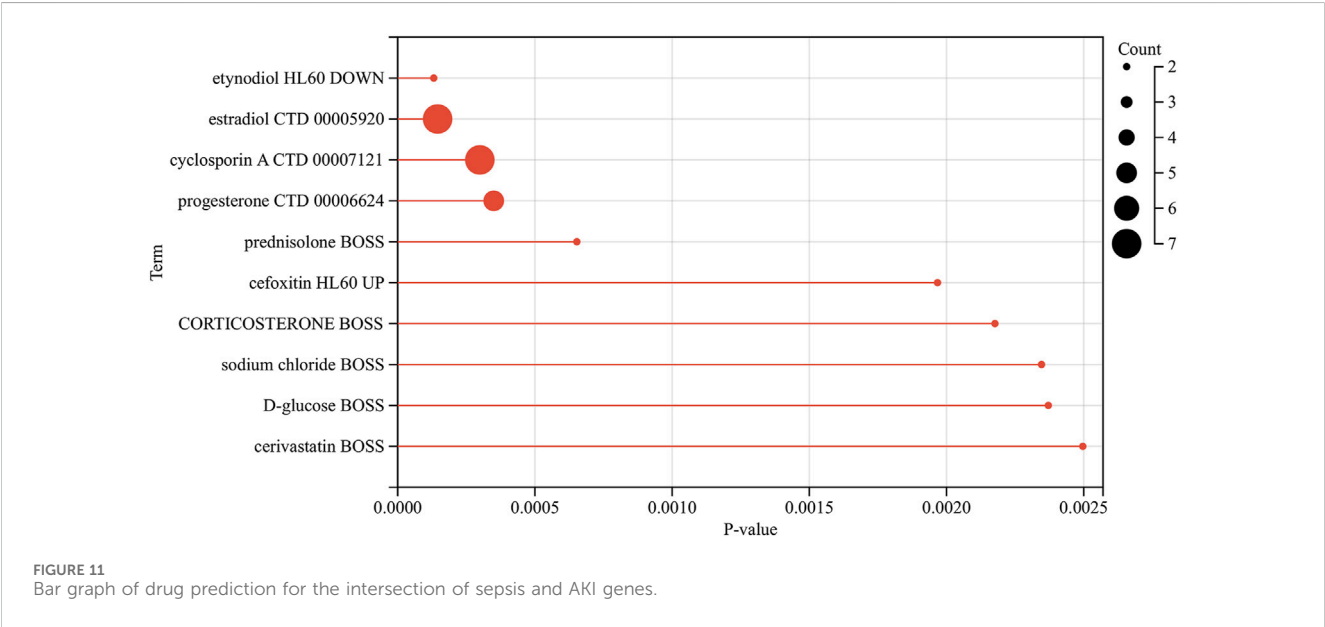


TABLE 2 Correlation of immune cells with model genes.

Data set	Gene name	Immune cell	Correlation coefficient
Sepsis	NR3C2	T.cells.CD4.naive	+0.70
Sepsis	NR3C2	Macrophages.M0	−0.60
Sepsis	PLEKHO1	Neutrophils	−0.61
Sepsis	PLEKHO1	Macrophages.M0	−0.65
Sepsis	CDC25B	Neutrophils	−0.72
Sepsis	CDC25B	Macrophages.M0	−0.60
Sepsis	CDC25B	NK cells.resting	+0.69
Sepsis	CDC25B	T.cells.CD4.naive	+0.63
Sepsis	VNN1	Neutrophils	+0.61
Sepsis	VNN1	Macrophages.M0	+0.69
Sepsis	VNN1	NK.cells.resting	−0.64
Sepsis	VNN1	T.cells.CD8	−0.71
Sepsis	SLC2A3	Neutrophils	+0.68
Sepsis	SLC2A3	Macrophages.M0	+0.63
Sepsis	SLC2A3	NK.cells.resting	−0.61
Sepsis	SLC2A3	T.cells.CD8	−0.70
Sepsis	RPL36	Neutrophils	−0.65
Sepsis	RPL36	T.cells.CD4.naive	+0.65
AKI	RPL36	Eosinophils	−0.63

AKI: acute kidney injury.



diagnostic accuracy based on the high AUC in both the internal and external validation sets, especially in the age under 60 years and male patient subgroups, which showed stronger diagnostic performance. At the same time, our model achieved diagnostic performance superior to that of the previous diagnostic model identified by [Tang et al. \(2021\)](#). that influences the composition of genes important for the development of SA-AKI, both in the training and validation cohorts.

In our model, four genes (*NR3C2*, *CEACAM1*, *VNN1*, and *SLC2A3*) were linked to the development of AKI in patients with sepsis. *NR3C2* participates in corticosteroid signaling, a process that is essential for modulating the inflammatory response observed in sepsis (Wong et al., 2016). The upregulation of *NR3C2* in the sepsis cohort suggests a compensatory mechanism designed to counteract the inflammatory milieu, which may have implications for therapeutic interventions targeting this receptor. *CEACAM1* is a member of the carcinoembryonic antigen family and plays a crucial role in immune regulation and cell-cell interactions. It modulates T cell responses and promotes inflammatory regression (Skubitz, 2024). In our AKI cohort, downregulation of *CEACAM1* may reflect impaired immune responses in the context of sepsis, leading to the development of SA-AKI. Downregulation of *SLC2A3* expression in sepsis decreases glucose uptake by kidney cells, leading to insufficient intracellular energy metabolism. The proper physiological functioning of renal cells is contingent upon a sufficient energy supply, and insufficient energy impairs tubular reabsorption, glomerular filtration and other functions, and triggers oxidative stress, which damages the biomolecules of kidney cells, further aggravating the kidney injury, and promoting the development of AKI (Luo et al., 2023). *VNN1* is involved in processes such as lipid metabolism (Ucakturk et al., 2016). *VNN1* triggers senescence in renal tubular cells by stimulating *RB1* expression (Chen et al., 2022). In animal models, elevated *VNN1* expression correlates with the degree of AKI and inhibition of *VNN1* activity attenuates pathological changes in the kidney (Hosohata et al., 2011). *PLEKHO1* is involved in a variety of cellular functions, including cytoskeletal arrangement and cell signaling (Nie et al., 2013). In this study, the different expression levels of *PLEKHO1* highlight its potential role as a biomarker and therapeutic target for SA-AKI. Few studies have assessed the direct involvement of *CDC25B*, *HEPACAM2*, and *RPL36* in the progression of SA-AKI. Although these genes have important functions in intracellular signaling, cell cycle, intercellular interactions, and protein synthesis, their specific mechanisms of action in SA-AKI have not been fully clarified. Consequently, additional investigations into their particular roles and mechanisms of action are essential to convert these discoveries into clinical applications and to avert the onset of SA-AKI.

Finally, through drug analysis on the Enrich platform, we identified 10 drug candidates, and the identification of potential therapeutic agents, including cyclosporine A and estradiol, opens new avenues for targeted therapeutic strategies. Cyclosporin A is a potent immunosuppressant that acts primarily on T lymphocytes (Flores et al., 2019). The primary mechanism may involve reducing the harm to the kidney tissue caused by the inflammatory response by suppressing the excessive activation of immune cells. Estradiol has antioxidant properties. It may have a therapeutic effect on AKI by blocking the mitochondrial apoptotic pathway through its direct effect on organelles (Borras et al., 2010), helping to maintain the normal structure and function of kidney cells. These drug candidates warrant further investigation through clinical trials to evaluate their efficacy in SA-AKI.

This study has some limitations. First, the lack of wet-laboratory validation limits the ability to confirm the biological relevance of

identified DEGs and their interactions. Although bioinformatics methods provide valuable insights, they cannot fully replace empirical data. Second, the longitudinal design of GSE57065 (with multiple time points per septic shock patient) may introduce intra-individual correlation bias, potentially inflating statistical significance for certain DEGs. Furthermore, while ComBat was applied to harmonize GSE57065 (septic shock) and GSE28750 (sepsis without shock), this batch correction method may inadvertently attenuate biologically meaningful time-dependent expression patterns. Third, despite the incorporation of multiple GEO datasets, the limited sample size could potentially restrict the broader applicability of the results obtained. Future studies should prioritize multi-omics validation in prospectively collected cohorts with standardized sampling time points while adopting advanced batch correction methods that better preserve temporal dynamics.

5 Conclusion

In conclusion, this study identified the key genes linking sepsis and AKI through comprehensive bioinformatics analysis. A model was developed for early AKI diagnosis in the context of sepsis using machine learning. These findings have the potential to advance the development of clinical diagnostic and co-treatment strategies, paving the way for improved prognosis of SA-AKI.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/geo/>, GSE57065 <https://www.ncbi.nlm.nih.gov/geo/>, GSE28750 <https://www.ncbi.nlm.nih.gov/geo/>, GSE30718 <https://www.ncbi.nlm.nih.gov/geo/>, GSE139061 <https://www.ncbi.nlm.nih.gov/geo/>, GSE67401.

Ethics statement

The studies involving humans were approved by Institutional Review Board of The Third People's Hospital Affiliated to Fujian University of Traditional Chinese Medicine. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants'; legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

CL: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing. MZ:

Formal Analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing. WW: Formal Analysis, Project administration, Software, Writing – review and editing. ZW: Formal Analysis, Methodology, Software, Writing – review and editing. GL: Project administration, Software, Supervision, Validation, Writing – review and editing. SF: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review and editing. XZ: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Fujian Provincial Health and Health Provincial Special Subsidy Funds (grant number: BPB - 2022FSD). The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, and approval of the manuscript; or the decision to submit the manuscript for publication.

Acknowledgments

We express our gratitude to GEO for offering the platform and to the data proprietors for supplying the data.

References

- Alsaeedi, A. H., Al-Mahmood, H. H. R., Alnaseri, Z. F., Aziz, M. R., Al-Shammary, D., Ibaida, A., et al. (2024). Fractal feature selection model for enhancing high-dimensional biological problems. *BMC Bioinforma.* 25, 12. doi:10.1186/s12859-023-05619-z
- Ayloo, S., Lazo, C. G., Sun, S., Zhang, W., Cui, B., and Gu, C. (2022). Pericyte-to-endothelial cell signaling via vitronectin-integrin regulates blood-CNS barrier. *Neuron* 110, 1641–1655.e6. doi:10.1016/j.neuron.2022.02.017
- Borras, C., Gambini, J., Lopez-Grueso, R., Pallardo, F. V., and Vina, J. (2010). Direct antioxidant and protective effect of estradiol on isolated mitochondria. *Biochim. Biophys. Acta* 1802, 205–211. doi:10.1016/j.bbadis.2009.09.007
- Brandstadter, J. D., and Yang, Y. (2011). Natural killer cell responses to viral infection. *J. Innate Immun.* 3, 274–279. doi:10.1159/000324176
- Chen, J., Lu, H., Wang, X., Yang, J., Luo, J., Wang, L., et al. (2022). VNN1 contributes to the acute kidney injury-chronic kidney disease transition by promoting cellular senescence via affecting RB1 expression. *FASEB J.* 36, e22472. doi:10.1096/fj.202200496RR
- Chen, W., Zhao, Y., Li, X. C., Kubiak, J. Z., Ghobrial, R. M., and Kloc, M. (2017). Rho-specific guanine nucleotide exchange factors (Rho-GEFs) inhibition affects macrophage phenotype and disrupts golgi complex. *Int. J. Biochem. Cell Biol.* 93, 12–24. doi:10.1016/j.biocel.2017.10.009
- Derive, M., and Gibot, S. (2011). Urine sTREM-1 assessment in diagnosing sepsis and sepsis-related acute kidney injury. *Crit. Care* 15, 1013. doi:10.1186/cc10533
- Enderlin Vaz Da Silva, Z., Lehr, H. A., and Velin, D. (2014). *In vitro* and *in vivo* repair activities of undifferentiated and classically and alternatively activated macrophages. *Pathobiology* 81, 86–93. doi:10.1159/000357306
- Ersavas, S., Diniz, G., Yildirim, H. T., Koca, Y., Kahraman, D. S., Ayaz, D., et al. (2016). Expression of neutrophil gelatinase-associated lipocalin and kidney injury Molecule-1 in Wilms tumor. *Turk Patoloji Derg.* 32, 158–163. doi:10.5146/tjpath.2015.01360
- Flores, C., Fouquet, G., Moura, I. C., Maciel, T. T., and Hermine, O. (2019). Lessons to learn from low-dose Cyclosporin-A: A new approach for unexpected clinical applications. *Front. Immunol.* 10, 588. doi:10.3389/fimmu.2019.00588
- Goggi, J. L., Qiu, L., Liao, M. C., Khanapur, S., Jiang, L., Boominathan, R., et al. (2020). Dopamine transporter neuroimaging accurately assesses the maturation of dopamine

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2025.1561331/full#supplementary-material>

neurons in a preclinical model of parkinson's disease. *Stem Cell Res. Ther.* 11, 347. doi:10.1186/s13287-020-01868-4

Grande, M. T., Perez-Barriocanal, F., and Lopez-Novoa, J. M. (2010). Role of inflammation in tubulo-interstitial damage associated to obstructive nephropathy. *J. Inflamm. (Lond)* 7, 19. doi:10.1186/1476-9255-7-19

Heung, M., and Chawla, L. S. (2014). Acute kidney injury: gateway to chronic kidney disease. *Nephron Clin. Pract.* 127, 30–34. doi:10.1159/000363675

Hosohata, K., Ando, H., Fujiwara, Y., and Fujimura, A. (2011). Vanin-1: a potential biomarker for nephrotoxicant-induced renal injury. *Toxicology* 290, 82–88. doi:10.1016/j.tox.2011.08.019

Hotchkiss, R. S., Tinsley, K. W., Hui, J. J., Chang, K. C., Swanson, P. E., Drewry, A. M., et al. (2000). p53-dependent and -independent pathways of apoptotic cell death in sepsis. *J. Immunol.* 164, 3675–3680. doi:10.4049/jimmunol.164.7.3675

Infinger, M., Grosse, J., Westphal, K., Leder, A., Ulbrich, C., Paul, M., et al. (2008). Vascular endothelial growth factor induces extracellular matrix proteins and osteopontin in the umbilical artery. *Ann. Vasc. Surg.* 22, 273–284. doi:10.1016/j.avsg.2007.11.002

Isali, I., McClellan, P., Wong, T. R., Cingireddi, S., Jain, M., Anderson, J. M., et al. (2022). *In vivo* delivery of M0, M1, and M2 macrophage subtypes via genipin-cross-linked collagen biotextile. *Tissue Eng. Part A* 28, 672–684. doi:10.1089/ten.TEA.2021.0203

Jensen, I. J., McGonagill, P. W., Butler, N. S., Harty, J. T., Griffith, T. S., and Badovinac, V. P. (2021). NK cell-derived IL-10 supports host survival during sepsis. *J. Immunol.* 206, 1171–1180. doi:10.4049/jimmunol.2001131

Jiang, W., and Zheng, R. (2022). Advances in the clinical value of tissue inhibitors of metalloproteinase-2 and insulin-like growth factor binding protein 7 in sepsis associated-acute kidney injury. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue* 34, 105–109. doi:10.3760/cma.j.cn121430-20210624-00123

Jung, J. W., Lee, A. R., Kim, J., Kim, Y. R., Jm, S. L., Lee, J. S., et al. (2021). Elucidating the functional roles of helper and cytotoxic T cells in the cell-mediated immune responses of olive flounder (*Paralichthys olivaceus*). *Int. J. Mol. Sci.* 22, 847. doi:10.3390/ijms22020847

Kellum, J. A., Wen, X., De Caestecker, M. P., and Hukriede, N. A. (2019). Sepsis-associated acute kidney injury: a problem deserving of new solutions. *Nephron* 143, 174–178. doi:10.1159/000500167

- Laurentino, M. R., Parente Filho, S. L. A., Parente, L. L. C., Da Silva Junior, G. B., Daher, E. F., and Lemes, R. P. G. (2019). Non-invasive urinary biomarkers of renal function in sickle cell disease: an overview. *Ann. Hematol.* 98, 2653–2660. doi:10.1007/s00277-019-03813-9
- Lee, H. Y., Kim, I. J., Choi, H. S., Jung, Y. H., Jeung, K. W., Mamadjonov, N., et al. (2024). Early single session of hyperbaric oxygen therapy mitigates renal apoptosis in lipopolysaccharides-induced endotoxemia in rats. *Kidney Res. Clin. Pract.* doi:10.23876/j.krcp.23.294
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883. doi:10.1093/bioinformatics/bts034
- Li, J., Zhu, M., and Yan, L. (2024). Predictive models of sepsis-associated acute kidney injury based on machine learning: a scoping review. *Ren. Fail* 46, 2380748. doi:10.1080/0886022X.2024.2380748
- Li, X., Wu, R., Zhao, W., Shi, R., Zhu, Y., Wang, Z., et al. (2023). Machine learning algorithm to predict mortality in critically ill patients with sepsis-associated acute kidney injury. *Sci. Rep.* 13, 5223. doi:10.1038/s41598-023-32160-z
- Lin, C., Wang, J., Cai, K., Luo, Y., Wu, W., Lin, S., et al. (2024). Elevated activated partial thromboplastin time as a predictor of 28-Day mortality in sepsis-associated acute kidney injury: a retrospective cohort analysis. *Int. J. Gen. Med.* 17, 1739–1753. doi:10.2147/IJGM.S459583
- Liu, Y., Sun, Q., Long, H., Qiu, Z., Zhang, D., Zhang, H., et al. (2024). The value of IL-6, PCT, qSOFA, NEWS, and SIRS to predict septic shock after percutaneous nephrolithotomy. *BMC Urol.* 24, 116. doi:10.1186/s12894-024-01502-y
- Liu, Z., Guo, C., Dang, Q., Wang, L., Liu, L., Weng, S., et al. (2022). Integrative analysis from multi-center studies identifies a consensus machine learning-derived lncRNA signature for stage II/III colorectal cancer. *EBioMedicine* 75, 103750. doi:10.1016/j.ebiom.2021.103750
- Luo, G., Wang, L., Zheng, Z., Gao, B., and Lei, C. (2023). Cuproptosis-related ferroptosis genes for predicting prognosis in kidney renal clear cell carcinoma. *Eur. J. Med. Res.* 28, 176. doi:10.1186/s40001-023-01137-z
- Lv, W., Liao, J., Li, C., Liu, D., Luo, X., Diao, R., et al. (2024). Aquaporin 1 is renoprotective in septic acute kidney injury by attenuating inflammation, apoptosis and fibrosis through inhibition of P53 expression. *Front. Immunol.* 15, 1443108. doi:10.3389/fimmu.2024.1443108
- Martensson, J., Bell, M., Oldner, A., Xu, S., Venge, P., and Martling, C. R. (2010). Neutrophil gelatinase-associated lipocalin in adult septic patients with and without acute kidney injury. *Intensive Care Med.* 36, 1333–1340. doi:10.1007/s00134-010-1887-4
- Martin, A. N., Alexander-Miller, M., Yoza, B. K., Vachharajani, V., and McCall, C. E. (2018). Sirtuin1 targeting reverses innate and adaptive immune tolerance in septic mice. *J. Immunol. Res.* 2018, 2402593. doi:10.1155/2018/2402593
- Montomoli, J., Donati, A., and Ince, C. (2019). Acute kidney injury and fluid resuscitation in septic patients: are we protecting the kidney? *Nephron* 143, 170–173. doi:10.1159/000501748
- Murugan, R., Karajala-Subramanyam, V., Lee, M., Yende, S., Kong, L., Carter, M., et al. (2010). Acute kidney injury in non-severe pneumonia is associated with an increased immune response and lower survival. *Kidney Int.* 77, 527–535. doi:10.1038/ki.2009.502
- Nie, J., Liu, L., He, F., Fu, X., Han, W., and Zhang, L. (2013). CKIP-1: a scaffold protein and potential therapeutic target integrating multiple signaling pathways and physiological functions. *Ageing Res. Rev.* 12, 276–281. doi:10.1016/j.arr.2012.07.002
- Odum, J. D., Standage, S., Alder, M., Zingarelli, B., Devarajan, P., and Wong, H. R. (2022). Candidate biomarkers for sepsis-associated acute kidney injury mechanistic studies. *Shock* 57, 687–693. doi:10.1097/SHK.0000000000001916
- Pillai, M., Adapa, K., Shumway, J. W., Dooley, J., Das, S. K., Chera, B. S., et al. (2022). Feature engineering for interpretable machine learning for quality assurance in radiation oncology. *Stud. Health Technol. Inf.* 290, 460–464. doi:10.3233/SHTI220118
- Qian, L., Sun, R., Aebersold, R., Buhlmann, P., Sander, C., and Guo, T. (2024). AI-empowered perturbation proteomics for complex biological systems. *Cell Genom* 4, 100691. doi:10.1016/j.xgen.2024.100691
- Qin, H., Abulaiti, A., Maimaiti, A., Abulaiti, Z., Fan, G., Aili, Y., et al. (2023). Integrated machine learning survival framework develops a prognostic model based on inter-crosstalk definition of mitochondrial function and cell death patterns in a large multicenter cohort for lower-grade glioma. *J. Transl. Med.* 21, 588. doi:10.1186/s12967-023-04468-x
- Reddy, S., Hu, B., and Kashani, K. (2023). Relationship between the rate of fluid resuscitation and acute kidney injury: a retrospective cohort study. *Int. J. Crit. Illn. Inj. Sci.* 13, 104–110. doi:10.4103/ijciis.ijciis_7_23
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-Sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007
- Skube, S. J., Katz, S. A., Chipman, J. G., and Tignanelli, C. J. (2018). Acute kidney injury and sepsis. *Surg. Infect. (Larchmt)* 19, 216–224. doi:10.1089/sur.2017.261
- Skubitz, K. M. (2024). The role of CEACAMs in neutrophil function. *Eur. J. Clin. Invest* 54 (Suppl. 2), e14349. doi:10.1111/eci.14349
- Souza, D. C., Barreira, E. R., Shieh, H. H., Ventura, A. M. C., Bousso, A., and Troster, E. J. (2021). Prevalence and outcomes of sepsis in children admitted to public and private hospitals in Latin America: a multicenter observational study. *Rev. Bras. Ter. Intensiva* 33, 231–242. doi:10.5935/0103-507X.20210030
- Su, Y., Gong, Z., Wu, Y., Tian, Y., and Liao, X. (2017). Diagnostic value of urine tissue inhibitor of Metalloproteinase-2 and insulin-like growth factor-binding protein 7 for acute kidney injury: a meta-analysis. *PLoS One* 12, e0170214. doi:10.1371/journal.pone.0170214
- Tang, Y., Yang, X., Shu, H., Yu, Y., Pan, S., Xu, J., et al. (2021). Bioinformatic analysis identifies potential biomarkers and therapeutic targets of septic-shock-associated acute kidney injury. *Hereditas* 158, 13. doi:10.1186/s41065-021-00176-y
- Ucakturk, A., Avci, B., Genc, G., Ozkaya, O., and Aydin, M. (2016). Kidney injury molecule-1 and neutrophil gelatinase associated lipocalin in normoalbuminuric diabetic children. *J. Pediatr. Endocrinol. Metab.* 29, 145–151. doi:10.1515/jpem-2015-0138
- Vodovar, D., Peyre, H., and Megarbane, B. (2021). Relationship between acute kidney injury and mortality in poisoning - a systematic review and meta-analysis. *Clin. Toxicol. (Phila)* 59, 771–779. doi:10.1080/15563650.2021.1928161
- Wei, X., Tu, Y., Bu, S., Guo, G., Wang, H., and Wang, Z. (2024). Unraveling the intricate web: complement activation shapes the pathogenesis of sepsis-induced coagulopathy. *J. Innate Immun.* 16, 337–353. doi:10.1159/000539502
- Wen, X., Li, S., Frank, A., Chen, X., Emlet, D., Hukriede, N. A., et al. (2020). Time-dependent effects of histone deacetylase inhibition in sepsis-associated acute kidney injury. *Intensive Care Med. Exp.* 8, 9. doi:10.1186/s40635-020-0297-3
- Wong, H. R., Atkinson, S. J., Cvijanovich, N. Z., Anas, N., Allen, G. L., Thomas, N. J., et al. (2016). Combining prognostic and predictive enrichment strategies to identify children with septic shock responsive to corticosteroids. *Crit. Care Med.* 44, e1000–e1003. doi:10.1097/CCM.0000000000001833
- Wu, J., and Lanier, L. L. (2003). Natural killer cells and cancer. *Adv. Cancer Res.* 90, 127–156. doi:10.1016/s0065-230x(03)90004-2
- Zarjou, A., and Agarwal, A. (2011). Sepsis and acute kidney injury. *J. Am. Soc. Nephrol.* 22, 999–1006. doi:10.1681/ASN.2010050484
- Zhang, J., Qi, M., Ma, L., and Liu, D. (2024). Research progress on biomarkers of sepsis-associated acute kidney injury. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue* 36, 1216–1220. doi:10.3760/cma.j.cn121430-20231018-00883