Machine learning-driven identification of the gene-expression signature associated with a persistent multiple organ dysfunction trajectory in critical illness

Mihir R. Atreya,[a,b,*,k] Shayantan Banerjee,[c,k] Andrew J. Lautz,[a,b] Matthew N. Alder,[a,b] Brian M. Varisco,[a,b] Hector R. Wong,[a,b,j] Jennifer A. Muszynski,[d,e] Mark W. Hall,[d,e] L. Nelson Sanchez-Pinto,[f,g] and Rishikesan Kamaleswaran,[h,i] for the Genomics of Pediatric Septic Shock Investigators

aDivision of Critical Care Medicine, Cincinnati Children's Hospital Medical Center and Cincinnati Children's Research Foundation, Cincinnati, 45229, OH, USA
bDepartment of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, 45267, USA
cDepartment of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai, 600 036, India
dDivision of Critical Care Medicine, Nationwide Children's Hospital, Columbus, 43205, OH, USA
eDepartment of Pediatrics, Ohio State University, Columbus, 43205, OH, USA
fDepartment of Pediatrics, Northwestern University Feinberg School of Medicine, Chicago, 60611, IL, USA
gDepartment of Health and Biomedical Informatics, Northwestern University Feinberg School of Medicine, Chicago, 60611, IL, USA
hDepartment of Biomedical Informatics, Emory University School of Medicine, Atlanta, 30322, GA, United States
iDepartment of Biomedical Engineering, Georgia Institute of Technology, Atlanta, 30322, GA, United States

Summary

Background Multiple organ dysfunction syndrome (MODS) disproportionately drives morbidity and mortality among critically ill patients. However, we lack a comprehensive understanding of its pathobiology. Identification of genes associated with a persistent MODS trajectory may shed light on underlying biology and allow for accurate prediction of those at-risk.

Methods Secondary analyses of publicly available gene-expression datasets. Supervised machine learning (ML) was used to identify a parsimonious set of genes associated with a persistent MODS trajectory in a training set of pediatric septic shock. We optimized model parameters and tested risk-prediction capabilities in independent validation and test datasets, respectively. We compared model performance relative to an established gene-set predictive of sepsis mortality.

Findings Patients with a persistent MODS trajectory had 568 differentially expressed genes and characterized by a dysregulated innate immune response. Supervised ML identified 111 genes associated with the outcome of interest on repeated cross-validation, with an AUROC of 0.87 (95% CI: 0.85–0.88) in the training set. The optimized model, limited to 20 genes, achieved AUROCs ranging from 0.74 to 0.79 in the validation and test sets to predict those with persistent MODS, regardless of host age and cause of organ dysfunction. Our classifier demonstrated reproducibility in identifying those with persistent MODS in comparison with a published gene-set predictive of sepsis mortality.

Interpretation We demonstrate the utility of supervised ML driven identification of the genes associated with persistent MODS. Pending validation in enriched cohorts with a high burden of organ dysfunction, such an approach may inform targeted delivery of interventions among at-risk patients.

---

*Corresponding author. Cincinnati Children's Hospital Medical Center, Division of Critical Care Medicine, MLC2005, 3333 Burnet Avenue, Cincinnati, OH, 45229, USA.
    E-mail address: Mihir.Atreya@cchmc.org (M.R. Atreya).
jDeceased.
kM.R.A. and S.B contributed equally to the work detailed in this manuscript.

Research in context

Evidence before this study
Numerous studies have focused on sepsis-related gene-expression profiles to predict mortality. However, few studies have explicitly sought to determine signatures associated with persistent multiple organ dysfunction syndrome (MODS)—a condition associated with high morbidity.

Added value of this study
We leveraged publicly available datasets to identify a gene-expression signature associated with a persistent MODS trajectory. By leveraging supervised machine learning, we identified a parsimonious set of genes and developed a classifier to reliably predict those with persistent MODS

among critically ill children and adults with different causes of organ dysfunction. Notably, our classifier had better reproducibility in identifying patients with the outcome of interest in test datasets, in comparison with an established gene set predictive of sepsis mortality.

Implications of all the available evidence
Our study highlights the utility of supervised machine learning to identify genes linked with a persistent MODS trajectory. Pending validation, our approach may facilitate early identification of at-risk patients and, in the future, may inform targeted delivery of interventions among those most critically ill.

Introduction
Multiple organ dysfunction syndrome (MODS) is a major cause for mortality among critically ill patients admitted to intensive care units (ICU).[1] Those who survive the acute phase remain at high-risk of new morbidity including technology dependence,[2] nosocomial infections,[3] and late death.[4,5] Despite the significant burden of disease, care for patients with MODS remains limited to organ support, with no disease modifying therapies consistently proven to improve clinical outcomes. Although numerous clinical phenotypes of sepsis associated MODS have been described,[3,6,7] their overlapping nature and the fact that a majority of patients cannot be assigned into any phenotype, makes discovery and delivery of appropriate therapeutic interventions challenging. Thus, approaches which facilitate a systematic understanding of MODS pathobiology and early identification of at-risk patients are the need of the hour.

Over the previous two decades, numerous studies have evaluated gene-expression profiles among critically ill patients to discover a signature associated with sepsis mortality.[8–11] Several have led to identification of genes and related protein biomarkers that have been useful to predict those at highest risk of poor outcomes.[12,13] Further, unsupervised clustering of these genes have been used to determine subclasses or 'endotypes' of sepsis,[8–10,14] of which those with a dysregulated adaptive immune response have demonstrated differential responses to receipt of corticosteroids.[15,16] Pending validation, such predictive enrichment strategies may be used to tailor therapies for patients with sepsis. It is conceivable that similar high throughput approaches, which shed light on MODS pathobiology, hold potential to facilitate development of precision medicine approaches to enhance patient care.

Few transcriptomic studies, however, have explicitly focused on MODS as the primary outcome.[17–20] Given the dynamic nature of critical illness and substantial morbidity associated with persistent MODS, it is conceivable that focusing on this subset of patients may facilitate early identification of at-risk patients. In the current study, we leveraged publicly available datasets to identify the gene-signatures associated with a persistent MODS trajectory among critically ill patients and unravelled biological mechanisms at play. We implemented supervised machine learning (ML) approaches to identify a parsimonious set of genes predictive of the outcome of interest, trained and validated a model to reliably identify those at high risk of persistent MODS, and demonstrated reproducibility of our approach across test datasets agnostic of host developmental age[21] and cause of organ dysfunctions. Finally, we tested whether our limited set of genes identified by us improved upon previously published gene sets demonstrated to predict sepsis mortality, in identifying those at risk of persistent MODS.

Methods
Ethics approval and consent to participate
Only de-identified clinical data and publicly available datasets were used for the conduct of this study. The characteristics of the datasets including details of RNA samples, the various time points of collection, and platforms used for gene expression across the datasets are detailed in Supplementary Table S1.

Training dataset
Microarray dataset GSE66099 obtained from paediatric septic shock patients[22] was downloaded from the NCBI Gene Expression Omnibus (GEO) repository. In this

dataset, data on organ dysfunction based on Proulx Criteria[23] were available between day 1 through 7 of paediatric intensive care unit (PICU) admission.[24] The primary comparison of interest was persistent MODS (death by day 7, persistence of ≥2 organ dysfunctions on day 7, or new MODS between days 1–7; n = 46), relative to a composite of those with resolving MODS (≥2 organ dysfunctions on day 1 or 3 and with <2 dysfunctions by day 3 and 7 respectively; n = 63) and those with no MODS (n = 92). The latter comprised of septic patients, non-septic patients with systemic inflammatory response syndrome (SIRS), and healthy controls. Our choice for this comparison was guided by the fact that patients with persistent MODS, despite intensive organ support, likely represent a subset of patients who may benefit from targeted therapeutic interventions, based on their underlying biological predisposition.

Differential expression of genes

The Affymetrix probes in the training dataset were matched to gene symbols using the Affymetrix Human Genome U133 Plus 2.0 (hgu133plus2.db). Data were pre-processed including batch correction for year of study and are detailed in Supplementary Tables S2 and S3, and Supplementary Figure S1. Differential expression of genes (DEGs) based on a log2 fold change ≥ ± 0.5, adjusted value for Benjamini Hochberg correction for false discovery rate <0.05, was performed using the limma package in R.[25] We conducted sensitivity analyses with and without inclusion of patients who died within the first 7 days, to test the premise that non-survivors may have a different signature relative to survivors with persistent organ dysfunctions. We used clusterProfiler[26] for functional gene enrichment, and CIBERSORT[27] a computational tool to deconvolute bulk expression data and estimate abundance of various immune cells subpopulations.

Supervised machine learning

The analytic approach across the various phases of the study is summarized in Fig. 1 and detailed in the Supplementary methods. **(1) Stratified 5-fold cross-validation**: Due to the skewed class distribution in the training dataset, we applied a 5-fold stratified cross-validation process, similar to those previously published by our group,[18] that involved randomly partitioning the training dataset into five equal subsets, with the class distribution of the dataset being preserved in each of the train-test splits. **(2) Feature selection**: In addition to genes identified by DEG analyses, we sought to apply novel analytic pipelines, including non-linear approaches, to identify an optimal set of highly discriminative genes. Due to the high dimensionality of the training dataset (n~22,000 genes), we sought to extract genes to distinguish patients with persistent MODS trajectory, relative to those with resolving or no MODS. We used three popular variable selection techniques including Least Absolute Shrinkage and Selection Operator (LASSO), Minimum Redundancy and Maximum Relevance (MRMR), and Random forests (RF) based variable importance technique. The genes selected by each of the above methods were aggregated into a single input feature set, and the list of DEGs obtained from the training dataset were added to this list. Subsequently, we deployed recursive feature elimination to reduce redundancy in genes identified through the various feature selection methods. **(3) Model fitting:** To counter the class imbalances in our training data, we incorporated both undersampling and oversampling techniques in the training dataset. Briefly, three binary classifications algorithms were used including logistic regression and two tree-based classifiers (Random Forest and Extra Trees classifiers). Hyper-parameter tuning was done using a cross-validated grid search technique on a subset of the training data over a parameter grid using the area under the curve as the scoring function. The Grid-SearchCV function with the default 3-fold cross-validation strategy and area under the receiver operating characteristic curve (AUROC) was used to search for the best set of hyperparameters.

To evaluate robustness of the model training and to ensure complete cross-validation, this entire process was repeated seven times, resulting in thirty-five unique train- and test-splits. The performances obtained during each run were averaged, and the summary AUROC along with 95% CI were reported. The fraction of times a particular gene was chosen out of the 35 train- and test-split runs was used to rank the genes in descending order of strength of association. Genes associated with the outcome of interest in ≥80% of repeated cross-validation experiments, an arbitrarily selected threshold, were selected for downstream analyses and model optimization.

Validation dataset

We used the validation E-MTAB-10938 ArrayExpress dataset published by *Snyder* et al. that consisted of 32 paediatric patients with septic shock, of some of whom had an immunoparalysis phenotype[19] for parameter tuning. Only 5 patients in this dataset met criteria for persistent MODS based on Proulx Criteria with the remaining 27 patients serving as comparison group. **Classification Model Parameter tuning**: We trained different feature sets (of sizes n = 5,10,15,..) of the top genes identified through the training dataset to tune the following parameters using the validation dataset: (1) optimal number of features, 2) sampling technique–classifier combination, and 3) optimal probability threshold, all of which are detailed in the Supplementary methods. Given the imbalanced classification problem, we experimented with different classification thresholds from 0 to 1 with step sizes of 0.001 and reported
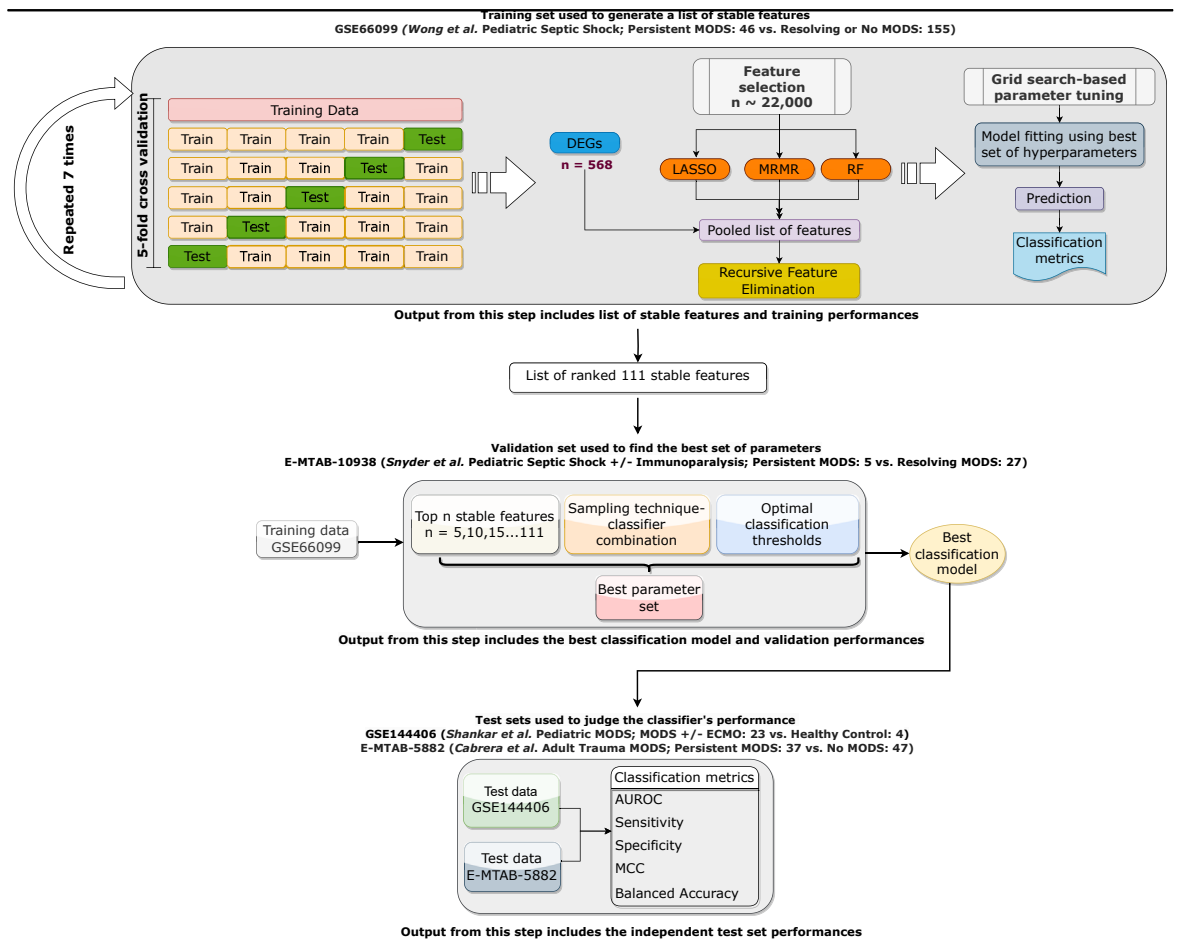
Fig. 1: Overarching approach of analyses and number of patients by outcome of interest in the training (GSE66099), validation (E-MTAB-10938), and two independent test sets (GSE144406 and E-MTAB-5882). Supervised machine learning approaches deployed in study including stratified 5-fold validation, feature selection, and model fitting phases are summarized. Differentially expressed genes (DEGs) identified in the training dataset were incorporated along with the pooled output of 3 feature selection methods including Least Absolute Shrinkage and Selection Operator (LASSO), Minimum Redundancy and Maximum Relevance (MRMR), and Random Forest (RF) based variable importance technique. Subsequently, we used recursive feature elimination and tested model performance to predict patients with persistent MODS trajectory. This entire process was repeated 7 times resulting in 35 train- and test-splits. The top 111 genes associated with outcome of interest in ≥80% of cross-validation experiments were selected for optimization in the validation dataset. We identified that top 20 genes and a fixed classification model had reproducible performance across 2 independent test datasets to predict persistent MODS.

model performance metrics at the threshold which provided the maximum AUROC.[28]

Test datasets
We tested the performance of the final model in two independent test datasets: (1) GSE144406 GEO dataset published by *Shankar* et al. that consisted of whole blood bulk RNA sequencing total of 23 paediatric patients with MODS, of whom six patients required extracorporeal membrane oxygenation (ECMO) support, and included four healthy controls[17] and (2) E-MTAB-5882 ArrayExpress dataset published by *Cabrera* et al. that consisted of time-course-based gene-expression profiling measurements collected from the whole blood of 70 critically injured adult patients in the hyper-acute time period within 2 hours of trauma.[29]

Model performance
Classification performance of models in the validation and test sets were judged based on the sensitivity, specificity, Matthew's Correlation Coefficient (MCC)—a balanced statistical measure of true positive, true negative, false positive, and false negatives,[30] and balanced accuracy—the arithmetic mean of sensitivity and specificity, in addition to AUROC. Model performance was reported at a fixed sensitivity of 85% across the validation and test datasets for ease of comparison of other classification metrics across datasets. The 95% CI for

each classification metric was derived by repeated sampling with replacement with 1000 iterations. The ci function from the gmodels package in R was used to calculate the CIs.

Performance relative to established genes predictive of sepsis mortality

We sought to determine whether genes identified through our supervised ML model were comparable or improved upon published literature on gene sets, which have been demonstrated to predict sepsis mortality, in identifying patients with persistent MODS. A total of 58 genes were outlined in *Sweeney* et al. that were predictive of 30-day mortality.[12] However, only 52/58 genes were present among validation and test datasets and were chosen for further analysis. We followed the same optimization as detailed above in the validation dataset but using 52 genes predictive of mortality, instead of the genes predictive of MODS identified by us.

Statistical analysis

Demographic and clinical data in the training dataset were summarized with counts and percentages or medians with interquartile ranges (IQR). Differences between groups were determined by χ^2 test for categorical variables and by one-way analysis of variance (ANOVA) for continuous variables. A p-value of 0.05 was used to test statistical significance, unless otherwise specified.

Role of funders

The content of the manuscript represents those of the authors and does not necessarily represent the official views of the National Institutes of Health (NIH). The funding agency did not have any role in study design, data collection, data analyses, interpretation, or writing of this report.

Results

A total of 201 patients with phenotyping of organ dysfunction trajectories were included in the training dataset. The demographic characteristics of the cohort are shown in Table 1. Forty-six patients had persistent MODS, including 15 patients who died within 7 days of study enrollment. Sixty-three patients had resolving MODS. Those with no MODS included 19 patients with sepsis without shock or organ dysfunctions on day 1, 26 patients with SIRS, and 47 patients admitted for elective surgical procedures who served as healthy controls. Patients with persistent MODS were younger, had higher illness severity at baseline, and a trend toward higher day 1 vasoactive inotropic scores (VIS). Unsurprisingly, those with persistent MODS trajectory had significantly higher 28-day mortality, fewer PICU free days, and higher cardiovascular, respiratory, and renal support requirements than those with resolving or no MODS. Individual organ dysfunctions and supportive interventions by MODS trajectory are detailed by day of septic shock in Supplementary Tables S4 and S5 respectively.

Gene-expression signature associated with persistent MODS trajectory and its biological relevance

568 genes were differentially expressed among patients with persistent MODS relative to those with resolving or no MODS; 369 genes were overexpressed, and 199 genes were underexpressed. The heat map and volcano plot for DEG analyses are shown in Fig. 2. In sensitivity analyses, exclusion of patients who died within the first 7 days did not significantly alter our results (data not shown). Pathway analyses revealed enrichment of neutrophil and cytokine signalling and cellular components involved in cytoplasmic and secretory granule lumen formation, as shown in Fig. 3. CIBERSORT analyses revealed that although neutrophils and monocytes accounted for the most abundant cell types, there were no significant differences among estimated cell proportions among those with persistent MODS trajectory relative to those without. However, an over-representation of M0 macrophages and plasma cells and an under-representation of CD8+ T cells, γδ T cells, and memory B cells was observed among patients with persistent MODS relative to those without, as shown in Fig. 4.

Genes associated with persistent MODS can be used to reliably identify those at-risk among children and adults with different causes for organ dysfunctions

The number of features identified by each of the feature selection methods is shown in Supplementary Figure S2. There was limited overlap of genes identified by LASSO, MRMR, RF feature selection methods. Through recursive feature elimination, we identified 111 genes consistently associated with persistent MODS trajectory in >80% of repeated cross-validation experiments, detailed in Supplementary Table S6. The summary AUROC for the risk prediction model across repeated cross-validation experiments to distinguish patients with persistent MODS trajectory relative to those with resolving or no MODS in the training dataset was 0.87 (95% CI: 0.85–0.88) with an MCC of 0.64 (95% CI: 0.60–0.68). The model had a sensitivity of 94.0% (87–93%) and specificity of 79% (76–83%).

In the validation dataset (E-MTAB-10938), we identified that the optimal model to predict those at risk of persistent MODS was achieved by using only the top 20 out of 111 genes identified through the training dataset. The genes identified were *RETN, ADAMTS3, LDHA, LCN2, IL1R2, DDIT4, CEACAM8, MERTK, MPO, ARL4A, CDKN3, PRTN3, MTMR11, ANLN, IL1RAP, HLA-DMB, ZBTB16, NUSAP1, GGH, and MMP8.*

	Persistent mods	Resolving mods	no mods	p-value
N (%)	46 (22.7%)	63 (31.0%)	92 (46.3%)	
Age (Years)	1.8 (0.5, 4.5)	2.4 (1.1, 5.2)	2.9 (1.3, 6.1)	0.03
Sex, M	28 (60.8%)	35 (55.5%)	50 (53.2%)	0.69
Race				
White	29 (63.0%)	40 (63.5%)	N/A	0.83
Black	11 (23.9%)	18 (28.5%)		
Other	6 (13.1%)	5 (8.0%)		
PRISM-III	21 (15, 29)	14 (10, 18)	1 (0, 10)	0.01
Day 1 VIS score	20 (1, 55)	10 (1, 20)	0 (0, 0)	0.07
Source				
Pulmonary	9 (20.9%)	7 (11.2%)	2 (2.2%)	0.45
Extrapulmonary	23 (53.4%)	28 (44.4%)	7 (7.6%)	
None	14 (30.4%)	28 (44.4%)	83 (90.2%)	
Pathogen type				0.66
Gram positive	19 (59.4%)	15 (42.8%)	5 (55.5%)	
Gram negative	10 (31.0%)	15 (42.8%)	6 (66.6%)	
Viral	2 (6.6%)	4 (11.4%)	0 (0)	
Fungal	1 (3.0%)	1 (2.8%)	0 (0)	
Outcomes				
28-Day mortality	17 (36.9%)	1 (1.5%)	0 (0)	<0.01
PICU free days	12 (0, 19)	22 (17, 24)	23 (19, 25)	<0.01
PICU LOS	10 (3, 19)	6 (4, 11)	5 (3, 8)	0.02
Hospital LOS	19 (3, 33)	10 (8, 21)	9 (7, 14)	0.38
Steroid use	18 (39.2%)	15 (30.6%)	4 (4.4%)	0.38

PRISM III: Pediatric Risk of Mortality III score; VIS score: Vasoactive inotropic score; LOS: Length of stay; N/A: Not available.

Table 1: **Demographic and outcome data by MODS trajectory in the training dataset.**

Each of these genes were selected by 2 or more of the feature selection methods employed. The best parameter set for classification of patients with persistent MODS vs. those without, using these 20 genes, were determined to be Standard Scaler, Instance Hardness Threshold (IHT) sampling technique, and Extra Trees (ET) classifier at a threshold of 0.488. In this dataset that included paediatric septic shock patients with and without an immunoparalysis phenotype, the AUROC of the model to predict persistent MODS was 0.74 (95% CI: 0.73–0.75). Finally, the AUROCs to predict MODS using a fixed set of 20 genes and classification parameters were 0.79 (95% CI: 0.78–0.80) in GSE144406 dataset of paediatric patients some of whom received ECMO support and 0.78 (95% CI: 0.77–0.78) in E-MTAB-5882 among adult patients in the hyper-acute
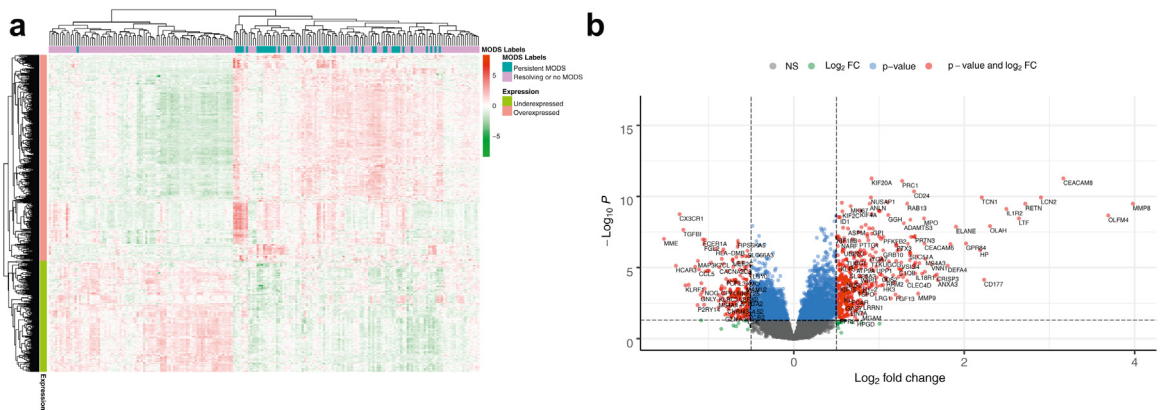
Fig. 2: a) Heat map, and b) Volcano plot of differentially expressed genes (DEGs) among patients with persistent MODS trajectory vs. those with resolving or no MODS in the training dataset.
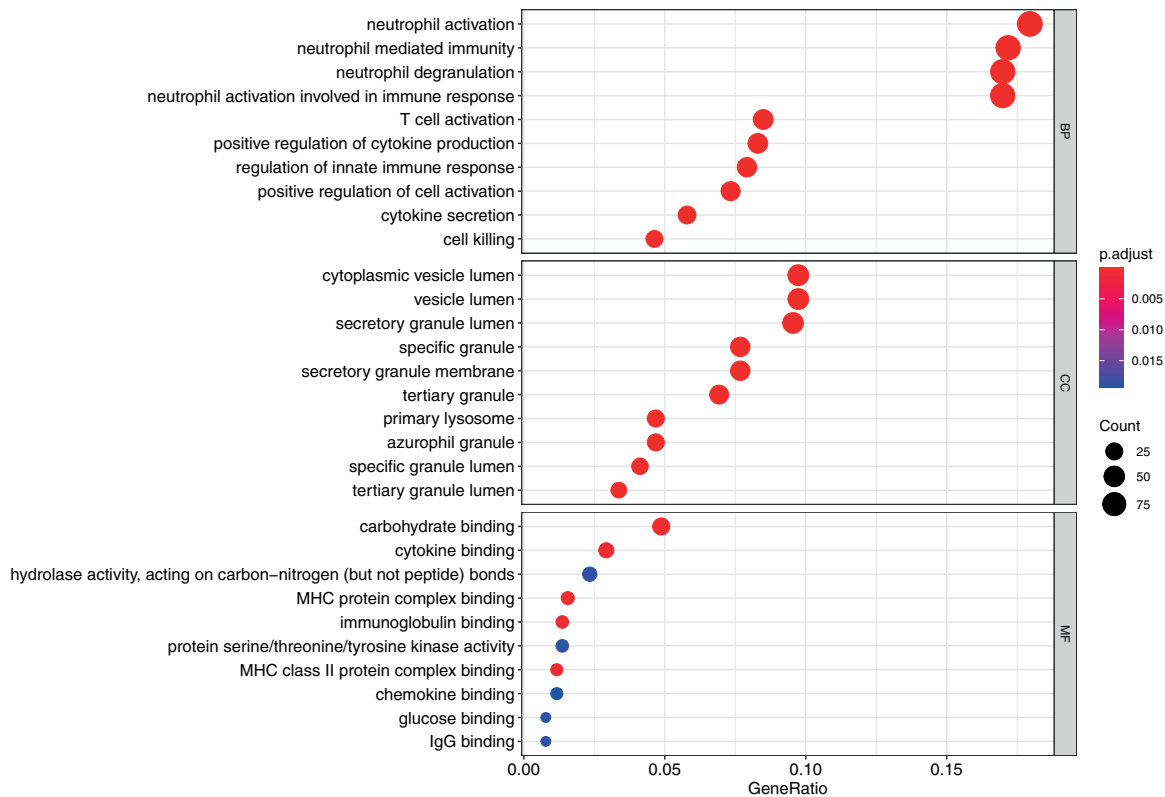
Fig. 3: Dot plot of biological pathways enriched among patients with a persistent MODS trajectory relative to those with resolving or no MODS. The size of the circles indicates the number of genes identified per pathway. The x-axis represents gene ratio which represents the fraction of differentially expressed genes found in the gene set. The color gradient shows the Benjamin Hochberg false discovery rate adjusted p-value for association between pathway and outcome of interest, with red indicating adj. p = ≤0.005. Abbreviations: BP: Biological process. CC: Cellular Components. MF: Molecular Functions.

phase of trauma. Model performance in validation and test sets are summarized in Table 2. The AUROCs of models in the training, validation and test sets are shown in Fig. 5.

Model performance compared to published gene sets predictive of sepsis mortality to predict risk of MODS

52 genes predictive of sepsis mortality with comparable ML model parameters including Extra Trees classifier model, MinMax Scaler, CC sampler, and a threshold of 0.429 were used to predict persistent MODS in the validation and test datasets. Although this gene-set demonstrated a comparable AUROC in the validation set (E-MTAB-10938), model performance in the test sets varied with AUROCs ranging from 0.57 to 0.76. These results are shown in Table 3.

Of note, 5 genes –*IL1R2, DDIT4, CEACAM8, MPO, and MTMR11*– overlapped between our 20 gene set predictive of persistent MODS trajectory and 52 gene set predictive of sepsis mortality published by Sweeney et al.[12] as highlighted in Supplementary Table S7.

Discussion
We present data demonstrating the gene-expression signature associated with a persistent MODS trajectory among critically ill patients, which was characterized by a dysregulated innate immune response. Further, by deploying supervised machine learning, we discovered a set of 111 genes consistently associated with a persistent MODS trajectory on repeated cross-validation experiments. Subsequently, we identified a parsimonious set of 20 genes and a fixed classifier model to reliably estimate risk of persistent MODS across validation and test datasets, including children and adults with different inciting causes for organ dysfunctions. Lastly, we demonstrate that our model had greater reproducibility in identifying patients with persistent MODS, relative to a gene-set previously established to predict sepsis mortality.

Gene-expression studies among paediatric patients with sepsis explicitly focused on MODS as an outcome have thus far been limited by patient sample size and case–control study design. Snyder et al. profiled 32 children with paediatric sepsis of whom 19
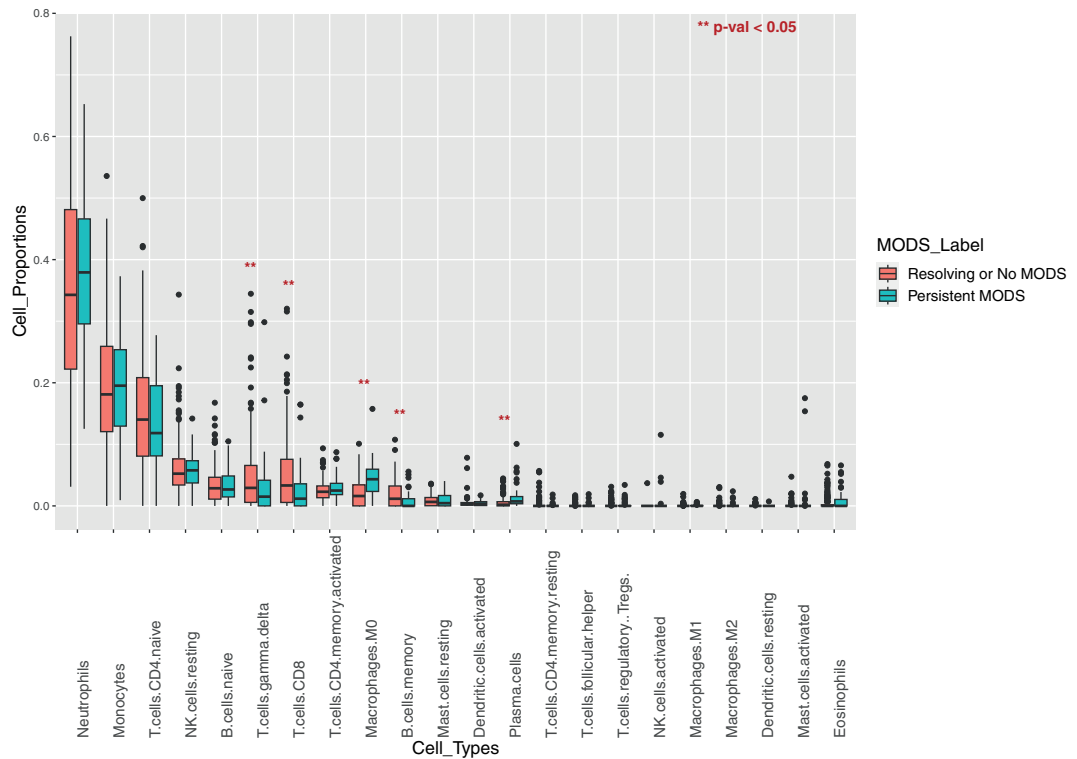
Fig. 4: Results of CIBERSORT analyses that show differences in proportions of major immune cell subsets among patients with persistent MODS trajectory vs. those with resolving or no MODS. Double asterisk denotes cell types with statistically significant differences between groups of interest.

had an immunoparalysis phenotype of MODS and identified 2303 DEGs, a majority of which were related to innate and adaptive immune systems.[19] Rama Shankar et al. used bulk RNA sequencing in a total of 27 paediatric septic shock patients and identified 30 DEGs when comparing those receiving ECMO support (n = 6) relative to those with MODS not receiving ECMO; a majority of genes belonged to the histone family.[17] In comparison, we used microarray data from a large prospective cohort of children with septic shock and identified 568 DEGs among patients with persistent MODS relative to those with resolving or no MODS.

The results of biological pathway analyses of gene-expression profiles associated with a persistent MODS trajectory demonstrated an overactive innate immune response with a key role for neutrophil degranulation. Although these results are unsurprising, there are several important considerations to be made. The gene-expression signature identified by us is very similar to prolonged MODS signature associated with paediatric patients with critical influenza, assessed by quantitative measurement of mRNA transcripts using a Nanostring platform.[20] In addition, they bear striking similarities with adults with a reactive or hyper-inflammatory high-risk phenotype of acute respiratory distress syndrome

Type	Dataset	Distribution	Sensitivity	Specificity	AUROC	Precision	MCC	Balanced accuracy
Validation	Pediatric Sepsis Immunoparalysis +/– (E-MTAB-10938)	#Positive: 5 #Negative: 27	0.85	0.70 (0.70–0.71)	0.75 (0.74–0.75)	0.33 (0.32–0.34)	0.36 (0.35–0.38)	0.775 (0.756–0.789)
Test	Pediatric Sepsis ECMO +/– (GSE144406)	#Positive: 4 #Negative: 23	0.85	0.30 (0.29–0.31)	0.79 (0.78–0.80)	0.94 (0.94–0.94)	0.13 (0.12–0.15)	0.575 (0.561–0.59)
Test	Adult Hyperacute phase of trauma (E-MTAB-5882)	#Positive: 37 #Negative: 47	0.85	0.51 (0.50–0.51)	0.78 (0.77–0.79)	0.58 (0.57–0.58)	0.36 (0.36–0.36)	0.679 (0.656–0.682)
Model parameters included top 20 genes, Standard Scaler, Instance Hardness Threshold sampling technique, and Extra-trees Classifier at a threshold of 0.488.								

Table 2: **Model performance across validation and test sets using 20 gene predictive of MODS and fixed parameters reported at a sensitivity of 85%.**
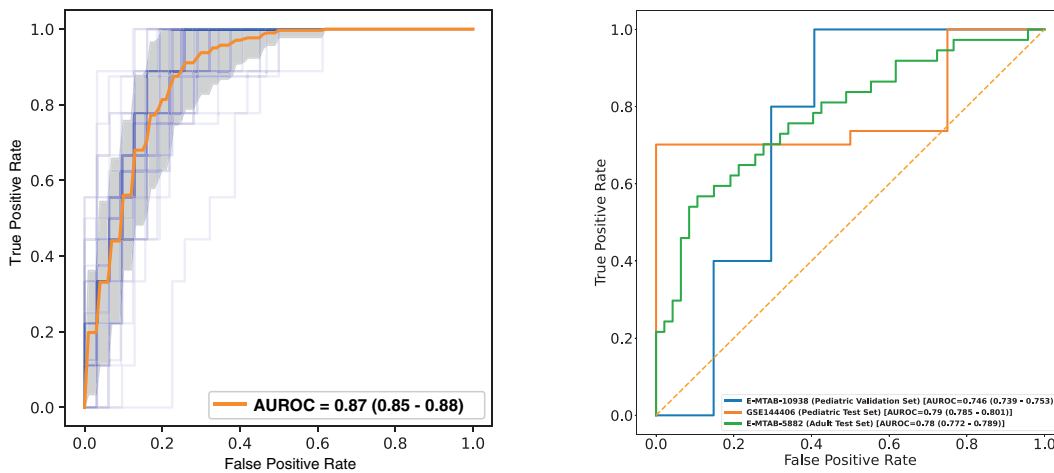
Fig. 5: Summary area under the receiver operating characteristic curve (AUROC) for the risk prediction model to estimate risk of persistent MODS in training dataset (GSE66099) across 35- train- and test-splits of repeated cross-validation experiments are shown in the left panel. The AUROCs of the final model which included top 20 genes identified through our study and an Extra Trees classifier model estimating risk of MODS across validation and two independent test sets is shown in the right panel.

(ARDS).[31] Among the top differentially expressed genes in our dataset, several including *RETN, LCN2, IL1R2, CEACAM8, and MPO* were all identified to contribute to neutrophil subset-specific responses and emergency granulopoiesis in multi-omic single cell analyses of immune cell subsets among septic patients by Kwok et al.[32] Taken together with the reproducibility of the predictive capabilities of our ML model to estimate risk of MODS across varying causes of organ dysfunctions including sepsis and trauma in the current study, we believe that our model is biologically relevant and may be generalizable across critical illness syndromes.

Results of CIBERSORT analyses revealed no significant differences in proportion of neutrophil or monocyte abundance between groups. However, we identified an overabundance of M0 macrophages and plasma cells and relatively fewer CD8+ T cells, γδ T cells, and memory B cells. This pattern of innate immune expansion with suppression of the adaptive immune arm has been consistently noted in sepsis,[12] and recently demonstrated to drive an extreme response endotype among septic patients.[32] Although our data are

extrapolations from bulk RNA microarrays, the consistency with single cell datasets strengthen our findings. Future research is necessary to determine the mechanistic link between immune cell subpopulations and organ dysfunction risk and whether targeted modulation of such immune cell subsets can be used as a novel therapeutic approach to promote organ recovery.

We demonstrate prognostic utility of our gene-expression classifier among critically ill patients with persistent organ dysfunctions. Our approach has several strengths including use of supervised ML to identify a limited set of genes consistently associated with the outcome of interest, model optimization in a separate validation set, and demonstration of reproducibility across two independent test sets. Our 20 gene-classifier more reliably predicted risk of persistent MODS in validation and test sets compared to an established 52 gene-set predictive of sepsis mortality, optimized through similar approaches. Our findings indicate that gene-expression signatures predictive of sepsis mortality may not be sufficient to consistently identify survivors with persistent organ failures nor generalizable across

Type	Dataset	Distribution	Sensitivity	Specificity	AUROC	Precision	MCC	Balanced accuracy
Validation	Pediatric Sepsis Immunoparalysis +/– (E-MTAB-10938)	#Positive: 5 #Negative: 27	0.85	0.70 (0.69–0.70)	0.76 (0.75–0.76)	0.33 (0.3–0.34)	0.37 (0.37–0.38)	0.775 (0.75–0.76)
Test	Pediatric Sepsis ECMO +/– (GSE144406)	#Positive: 4 #Negative: 23	0.85	0.0	0.57 (0.56–0.57)	0.92 (0.91–0.92)	−0.1 (−0.15- (−0.08))	0.425 (0.43–0.43)
Test	Adult Hyperacute phase of trauma (E-MTAB-5882)	#Positive: 37 #Negative: 47	0.85	0.34 (0.33–0.34)	0.76 (0.75–0.77)	0.507 (0.50–0.51)	0.23 (0.22–0.24)	0.595 (0.60–0.61)
Model parameters included 51 genes predictive of sepsis mortality, MinMaxScaler, CC sampling technique, and Extra-trees Classifier at a threshold of 0.429.								

Table 3: Model performance to estimate risk of MODS across validation and test sets using 51 gene predictive of sepsis mortality and fixed parameters reported at a sensitivity of 85%.

various phenotypes of MODS. In future studies, we will seek to prospectively validate our findings and leverage the gene-expression signature of persistent MODS to identify biologically relevant subclasses or endotypes, which may hold potential to demonstrate heterogeneity of treatment effect with modulators of the innate immune response among patients.[33–35]

Our data has several limitations. (1) The sample size of our training cohort and the percentage of patients who had persistent MODS trajectory were relatively limited. We therefore did not have sufficient statistical power to determine the signature of persistent MODS relative to those with a resolving MODS trajectory alone. Accordingly, we were forced to include those with no MODS in the comparison group, which may have contributed to diluting the signal of interest. (2) We used microarray data in the training dataset and identified a relatively small set of genes associated with organ dysfunction trajectories. Bulk RNA-sequencing is likely to provide a wider dynamic range including novel and low-abundance transcripts with higher sensitivity and specificity. (3) Although we identified the shared signature associated with MODS, it is likely that further heterogeneity exists. Large cohorts enriched for children and adults with MODS may shed further light on the underlying biology and address each of the above concerns. (4) Gene-expression data were collected at a single time point. However, temporal transcriptomic shifts and endotyping class switching are well documented between day 1 and 3 in paediatric septic shock.[36] Further, evolution of organ dysfunctions in sepsis is dynamic and influenced by both the underlying biology and the interventions used to support organs. Thus, sampling the transcriptome at multiple time points to determine gene-expression trajectories may better inform the evolution of organ dysfunctions among those critically[37] (5) We used only gene-expression datasets and did not sample other 'omic' layers. Multi-omics approaches may thus serve to deepen our understanding of MODS pathobiology and identify causal genes with mechanistic relevance.[38]

Conclusions
We provide evidence for a gene-expression signature associated with persistent MODS trajectory. Pending validation in enriched cohorts with a high burden of organ dysfunctions, our gene-expression classifier may facilitate the early identification of high-risk critically ill patients who may benefit from targeted therapies, including those that modulate the innate immune response.

Contributors
Study conceptualization: M.R.A., H.R.W., and R.K. Due to his untimely death, H.R.W. could not contribute to the drafting of the manuscript. Data analysis: M.R.A., S.B., and R.K have accessed, verified the underlying data, and conducted analyses detailed in the manuscript. Drafting of the manuscript: M.R.A and S.B. Critical review of the manuscript for important intellectual content: M.R.A, S.B., A.J.L, M.N.A, B.M.V, J.A.M, M.W.H, N.S.P, and R.K. All authors have approved the final version of this manuscript.

References
1 Weiss SL, Fitzgerald JC, Pappachan J, et al. Global epidemiology of pediatric severe sepsis: the sepsis prevalence, outcomes, and therapies study. *Am J Respir Crit Care Med*. 2015;191(10):1147–1157.
2 Carlton EF, Donnelly JP, Hensley MK, Cornell TT, Prescott HC. New medical device acquisition during pediatric severe sepsis hospitalizations. *Crit Care Med*. 2020;48(5):725–731.
3 Hall MW, Knatz NL, Vetterly C, et al. Immunoparalysis and nosocomial infection in children with multiple organ dysfunction syndrome. *Intensive Care Med*. 2011;37(3):525–532.
4 Zimmerman JJ, Banks R, Berg RA, et al. Critical illness factors associated with long-term mortality and Health related quality of life morbidity following community-acquired pediatric septic shock. *Crit Care Med*. 2020;48(3):319–328.
5 Wiens MO, Bone JN, Kumbakumba E, et al. Mortality after hospital discharge among children younger than 5 years admitted with suspected sepsis in Uganda: a prospective, multisite, observational cohort study. *Lancet Child Adolesc Health*. 2023;7(8):555–566.
6 Carcillo JA, Podd B, Aneja R, et al. Pathophysiology of pediatric multiple organ dysfunction syndrome. *Pediatr Crit Care Med*. 2017;18(3_suppl Suppl 1):S32–S45.
7 Carcillo JA, Halstead ES, Hall MW, et al. Three hypothetical inflammation pathobiology phenotypes and pediatric sepsis-induced multiple organ failure outcome. *Pediatr Crit Care Med*. 2017;18(6):513–523.
8 Wong HR, Cvijanovich N, Lin R, et al. Identification of pediatric septic shock subclasses based on genome-wide expression profiling. *BMC Med*. 2009;7:34.
9 Davenport EE, Burnham KL, Radhakrishnan J, et al. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. *Lancet Respir Med*. 2016;4(4):259–271.
10 Scicluna BP, van Vught LA, Zwinderman AH, et al. Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study. *Lancet Respir Med*. 2017;5(10):816–826.
11 Pelaia TM, Shojaei M, McLean AS. The role of transcriptomics in redefining critical illness. *Crit Care*. 2023;27(1):89.

12 Sweeney TE, Perumal TM, Henao R, et al. A community approach to mortality prediction in sepsis via gene expression analysis. *Nat Commun*. 2018;9(1):694.

13 Cano-Gamez E, Burnham KL, Goh C, et al. An immune dysfunction score for stratification of patients with acute infection based on whole-blood gene expression. *Sci Transl Med*. 2022;14(669):eabq4433.

14 Sweeney TE, Azad TD, Donato M, et al. Unsupervised analysis of transcriptomics in bacterial sepsis across multiple datasets reveals three robust clusters. *Crit Care Med*. 2018;46(6):915–925.

15 Wong HR, Cvijanovich NZ, Anas N, et al. Developing a clinically feasible personalized medicine approach to pediatric septic shock. *Am J Respir Crit Care Med*. 2015;191(3):309–315.

16 Antcliffe DB, Burnham KL, Al-Beidh F, et al. Transcriptomic signatures in sepsis and a differential response to steroids. From the VANISH randomized trial. *Am J Respir Crit Care Med*. 2019;199(8):980–986.

17 Shankar R, Leimanis ML, Newbury PA, et al. Gene expression signatures identify paediatric patients with multiple organ dysfunction who require advanced life support in the intensive care unit. *eBioMedicine*. 2020;62:103122.

18 Banerjee S, Mohammed A, Wong HR, Palaniyar N, Kamaleswaran R. Machine learning identifies complicated sepsis course and subsequent mortality based on 20 genes in peripheral blood immune cells at 24 H post-ICU admission. *Front Immunol*. 2021;12:592303.

19 Snyder A, Jedreski K, Fitch J, et al. Transcriptomic profiles in children with septic shock with or without immunoparalysis. *Front Immunol*. 2021;12:733834.

20 Novak T, Crawford JC, Hahn G, et al. Transcriptomic profiles of multiple organ dysfunction syndrome phenotypes in pediatric critical influenza. *Front Immunol*. 2023;14:1220028.

21 Wynn JL, Cvijanovich NZ, Allen GL, et al. The influence of developmental age on the early transcriptomic response of children with septic shock. *Mol Med*. 2011;17(11-2):1146–1156.

22 Sweeney TE, Shidham A, Wong HR, Khatri P. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Sci Transl Med*. 2015;7(287):287ra71.

23 Proulx F, Fayon M, Ann Farrell C, Lacroix J, Gauthier M. Epidemiology of sepsis and multiple organ dysfunction syndrome in children. *Chest*. 1996;109(4):1033–1037.

24 Atreya MR, Cvijanovich NZ, Fitzgerald JC, et al. Integrated PERSEVERE and endothelial biomarker risk model predicts death and persistent MODS in pediatric septic shock: a secondary analysis of a prospective observational study. *Crit Care*. 2022;26(1):210.

25 Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.

26 Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–287.

27 Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–457.

28 Zhou ZH, Liu XY. Training cost-sensitive neural networks with methods addressing the class imbalance problem. *IEEE Trans Knowl Data Eng*. 2006;18(1):63–77.

29 Cabrera CP, Manson J, Shepherd JM, et al. Signatures of inflammation and impending multiple organ dysfunction in the hyper-acute phase of trauma: a prospective cohort study. *PLoS Med*. 2017;14(7):e1002352.

30 Chicco D, Jurman G. An invitation to greater use of matthews correlation coefficient in robotics and artificial intelligence. *Front Robot AI*. 2022[cited 2022 Jul 6];9. Available from: https://www.frontiersin.org/articles/10.3389/frobt.2022.876814.

31 Bos LDJ, Scicluna BP, Ong DSY, Cremer O, van der Poll T, Schultz MJ. Understanding heterogeneity in biologic phenotypes of acute respiratory distress syndrome by leukocyte expression profiles. *Am J Respir Crit Care Med*. 2019;200(1):42–50.

32 Kwok AJ, Allcock A, Ferreira RC, et al. Neutrophils and emergency granulopoiesis drive immune suppression and an extreme response endotype during sepsis. *Nat Immunol*. 2023;24(5):767–779.

33 Bo L, Wang F, Zhu J, Li J, Deng X. Granulocyte-colony stimulating factor (G-CSF) and granulocyte-macrophage colony stimulating factor (GM-CSF) for sepsis: a meta-analysis. *Crit Care*. 2011;15(1):R58.

34 Interleukin-6 receptor antagonists in critically ill patients with covid-19. *N Engl J Med*. 2021;384(16):1491–1502.

35 Karakike E, Scicluna BP, Roumpoutsou M, et al. Effect of intravenous clarithromycin in patients with sepsis, respiratory and multiple organ dysfunction syndrome: a randomized clinical trial. *Crit Care*. 2022;26:183.

36 Wong HR, Cvijanovich NZ, Anas N, et al. Endotype transitions during the acute phase of pediatric septic shock reflect changing risk and treatment response. *Crit Care Med*. 2018;46(3):e242–e249.

37 Maslove DM, Wong HR. Gene expression profiling in sepsis: timing, tissue, and translational considerations. *Trends Mol Med*. 2014;20(4):204–213.

38 Chen T, Conroy J, Wang X, et al. The independent prognostic value of global epigenetic alterations: an analysis of single-cell ATAC-seq of circulating leukocytes from trauma patients followed by validation in whole blood leukocyte transcriptomes across three etiologies of critical illness. *eBioMedicine*. 2022;76:103860.