EBioMedicine 62 (2020) 103081

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Deep learning-based clustering robustly identified two classes of sepsis with both prognostic and predictive values

Zhongheng Zhang[a,*], Qing Pan[b], Huiqing Ge[c], Lifeng Xing[a], Yucai Hong[a], Pengpeng Chen[a]

[a] Department of Emergency Medicine, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, 310016, China
[b] College of Information Engineering, Zhejiang University of Technology, 310023, Hangzhou, China
[c] Department of Respiratory Care, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China

ARTICLE INFO

Article History:
Received 6 August 2020
Revised 19 September 2020
Accepted 7 October 2020
Available online xxx

Key words:
Sepsis
Endotype
Transcriptome
Autoencoder

ABSTRACT

Background: Sepsis is a heterogenous syndrome and individualized management strategy is the key to successful treatment. Genome wide expression profiling has been utilized for identifying subclasses of sepsis, but the clinical utility of these subclasses was limited because of the classification instability, and the lack of a robust class prediction model with extensive external validation. The study aimed to develop a parsimonious class model for the prediction of class membership and validate the model for its prognostic and predictive capability in external datasets.

Methods: The Gene Expression Omnibus (GEO) and ArrayExpress databases were searched from inception to April 2020. Datasets containing whole blood gene expression profiling in adult sepsis patients were included. Autoencoder was used to extract representative features for k-means clustering. Genetic algorithms (GA) were employed to derive a parsimonious 5-gene class prediction model. The class model was then applied to external datasets (n = 780) to evaluate its prognostic and predictive performance.

Findings: A total of 12 datasets involving 1613 patients were included. Two classes were identified in the discovery cohort (n = 685). Class 1 was characterized by immunosuppression with higher mortality than class 2 (21.8% [70/321] vs. 12.1% [44/364]; $p < 0.01$ for Chi-square test). A 5-gene class model (C14orf159, AKNA, PILRA, STOM and USP4) was developed with GA. In external validation cohorts, the 5-gene class model (AUC: 0.707; 95% CI: 0.664 − 0.750) performed better in predicting mortality than sepsis response signature (SRS) endotypes (AUC: 0.610; 95% CI: 0.521 − 0.700), and performed equivalently to the APACHE II score (AUC: 0.681; 95% CI: 0.595 − 0.767). In the dataset E-MTAB-7581, the use of hydrocortisone was associated with increased risk of mortality (OR: 3.15 [1.13, 8.82]; p = 0.029) in class 2. The effect was not statistically significant in class 1 (OR: 1.88 [0.70, 5.09]; p = 0.211).

Interpretation: Our study identified two classes of sepsis that showed different mortality rates and responses to hydrocortisone therapy. Class 1 was characterized by immunosuppression with higher mortality rate than class 2. We further developed a 5-gene class model to predict class membership.

Funding: The study was funded by the National Natural Science Foundation of China (Grant No. 81,901,929).

Introduction

Sepsis is a heterogenous syndrome consistingof many disease entities, all with individual pathophysiology and variations in host response [1–3]. Undifferentiated sepsis syndromes, therefore, lack genuine homogeneity and are the wrong point of entry for studies. That is why many large sepsis trials fail to identify a statistically and clinically significant results [4–7]. Thus, it is unwise to use the "one-size-fit-all" model in the management of critically ill patients with sepsis. Many efforts have been made to realize the individualized treatment strategy for sepsis [8, 9]. Our study group previously reported that critically ill patients with sepsis could be categorized into subphenotypes and these subphenotypes responded differently to the amount of fluid infusion (Z. [10]). Others also identified several subphenotypes by utilizing large clinical database and showed that these subphenotypes can have important implications for clinical treatment and trial design [3,11,12]. However, these studies input routinely collected clinical data as features for clustering analysis. While these features are readily available even in resource limited conditions, they may fail to capture important pathophysiological

* Corresponding to: Zhongheng Zhang, Address: No 3, East Qingchun Road, Hangzhou 310016, Zhejiang Province, China
E-mail addresses: zh_zhang1984@zju.edu.cn (Z. Zhang), pqpq@zjut.edu.cn (Q. Pan), netghq@zju.edu.cn (H. Ge), 3416231@zju.edu.cn (L. Xing), 407389157@qq.com (P. Chen).

https://doi.org/10.1016/j.ebiom.2020.103081

Research in Context

Evidence before this study

Sepsis is a heterogenous syndrome consisting of many disease entities, all with individual pathophysiology and variations in host response. Thus, individualized management strategy is the key to successful treatment. Genome wide expression profiling has been utilized for identifying subclasses of sepsis, but the clinical utility of these subclasses was limited because of the classification instability, and the lack of a robust class prediction model with extensive external validation.

Added value of this study

The study identified two endotypes of sepsis by using deep learning-based approach. The novel classification system identified two subtypes of sepsis that showed different responses to hydrocortisone therapy.

Implications of all the available evidence

The classification system developed in our study showed good prognostic and predictive values for the management of sepsis in clinical practice. Different subtypes of sepsis should be treated with different therapies. The functional enrichment analysis also revealed differing biological mechanisms underlying these subtypes.

alterations and cannot reveal underlying mechanisms. Pathophysiologically, thousands of genes are differentially expressed in response to infectious stimulus [13], and some of these genes can provide important prognostic and predictive information. We use the terms "prognostic" and "predictive" to refer to different things in the study. A prognostic biomarker/scores/model informs about a likely outcome (eg, disease recurrence, disease progression, death) independent of treatment received. A biomarker is predictive if the treatment effect (experimental compared with control) is different for biomarker-positive patients compared with biomarker-negative patients [14, 15].

Many studies have reported the genome-wide expression profiling of sepsis [13,16,17]. The differential gene expression has been well characterized between sepsis versus non-sepsis inflammatory responses [18], survivors versus non-survivors [19, 20], and sepsis of viral versus bacterial causes [21]. Two endotypes of sepsis were identified [20,22,23]. However, these studies developed prognostic models with complex machine learning algorithms involving tens of thousands of genes, making the models difficult to use. Furthermore, clustering methods were applied in a high-dimensional space, making the distances between points become relatively uniform, and the efficiency of clustering analysis could be reduced. Finally, studies combined children and adult population for pooled analysis [24]. We felt that more meaningful subclasses could be identified by focusing on more homogenous populations.

The availability of a large number of transcriptomic profiling databases provides unprecedented opportunity to use sophisticated deep learning algorithms for discovering novel prognostic and predictive biomarkers. These datasets also allow extensive external validations. Our study systematically reviewed available transcriptomic profiling datasets. We first used autoencoder to extract important representative features of the transcriptomic profile and then performed k-means clustering analysis on this lower dimension space. We further developed a parsimonious prediction model to predict the class membership obtained by clustering (class model). The class model was then used to predict class membership in external datasets. The predictive and prognostic values of the class membership were evaluated in several independent datasets. Interactions between class membership and use of hydrocortisone versus placebo, and vasopressin versus norepinephrine were explored. We hypothesized that the newly inferred classes of sepsis can have good prognostic and predictive performance. We have tried to overcome the limitation of overfitting by firstly performing dimension reduction with autoencoders; and then GA was employed for feature selection. Finally, we validated the model performance in external datasets (non-overlapping).

Methods

Datasets

The Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) and ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) databases were searched from inception to April 2020 to identify relevant transcriptomic profiling datasets. Datasets containing whole blood transcriptomic profiling in Homo sapiens were potentially eligible. Datasets were excluded if they included pediatric patients, were in vitro experiments, not assaying whole blood samples, not measuring mRNA, focusing on diseases caused by special pathogen such as hepatitis, human immunodeficiency virus (HIV) and hepatitis B virus (HBV). Furthermore, datasets did not report mortality outcome, or did not contain complete expression dataset were also excluded. Additional datasets could be added by manual search of the reference of included studies. The workflow of the study is shown in Fig. 1.

Ethics

The study utilized publicly available datasets with preexisting ethics approval from original studies. Informed consent was obtained for each participant. The study was conducted in accordance to the Helsinki declaration.

Datasets normalization and datasets combination

The raw datasets were downloaded and normalized according to the original studies. Probes with missing gene symbols were excluded. The maximum expression intensity was used when there were multiple probe sets mapping to the same gene symbol. The batch effects were adjusted for by using an empirical Bayes framework [25,26]. Common genes across all included datasets were retained for model discovery and validation.

Autoencoder

An autoencoder is a type of deep learning neural network used to learn efficient data representation in an unsupervised manner. The aim of an autoencoder is to learn a representation (encoding) for a set of data, typically for dimensionality reduction, by training the network to ignore signal "noise". The benefit of autoencoder against the commonly used principal component analysis is that it can perform non-linear dimensionality reduction [27]. Technical details of the training of autoencoder are shown in supplemental digital content (SDC Table S1).

As recommended by other studies [28,29], we reduced the original high-dimensional space into a space comprising 100 compressed representative features. These features were then filtered ($p < 0.05$) with univariate analysis against mortality outcome, aiming to ensure that the class membership obtained in subsequent k-mean clustering could have enough prognostic power.

k-means clustering analysis

We obtained 50 representative features after univariate filtering on mortality. These features were then used for k-means clustering

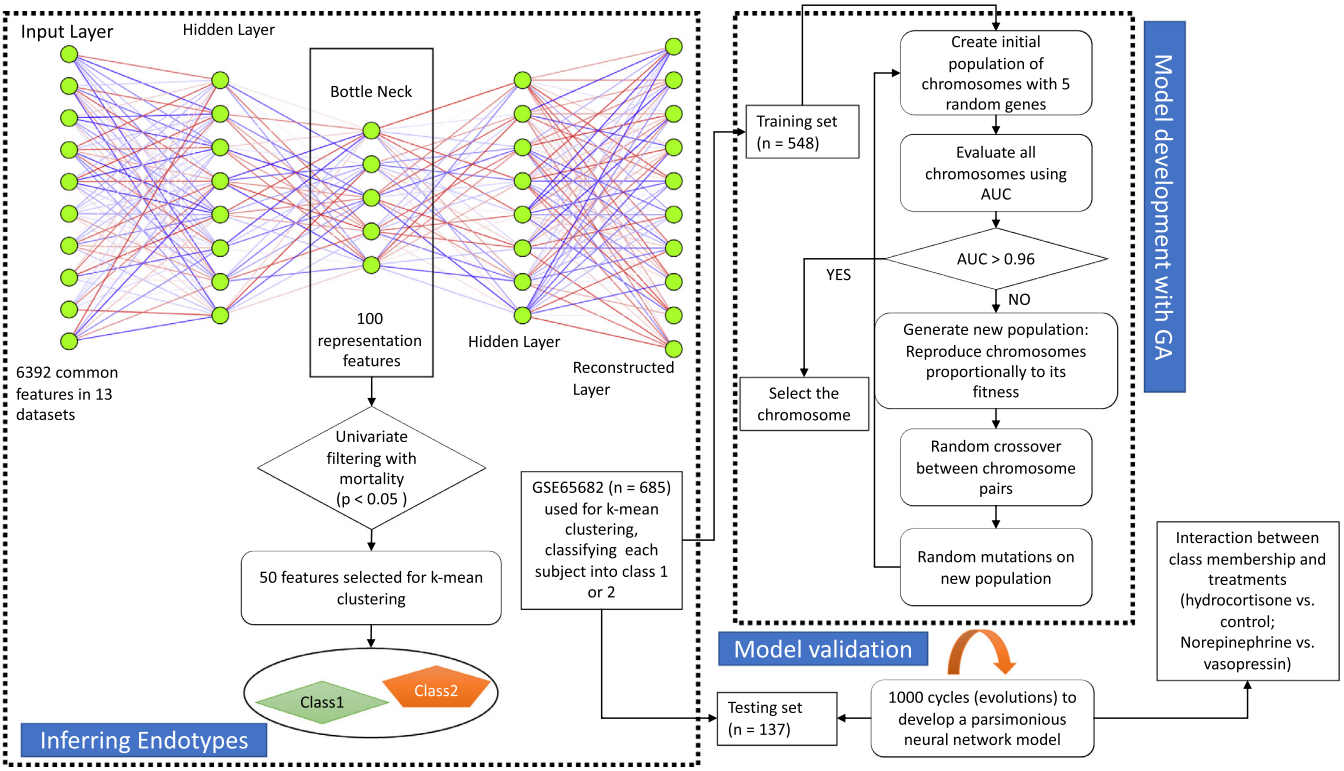

Fig. 1. Workflow of the study.

analysis. The number of clusters was determined by Elbow Method and Average Silhouette Method [30]. The number of classes was further confirmed by a Monte Carlo based approach for testing statistical significance [31]. The clustering analysis was performed in the discovery dataset (GSE65682).

Functional analysis of the two classes

Gene differential expression (GDE) was analysed between the two classes (class 1-class 2). Over- or under-expressed genes were analysed using the *limma* package and visualized with volcano plot [32]. Differentially expressed genes were subject to gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. Both gene set enrichment analysis (GSEA) and overrepresentation analysis were performed [33].

Genetic algorithms to develop a parsimonious model to predict class membership

A parsimonious class prediction model was developed for the ease of clinical utility [34]. The discovery cohort (GSE65682) was split into training and testing subsets by the ratio of 4:1. The training set (n = 548) was used to develop a model for the prediction of class membership derived from above mentioned k-means clustering, and the prediction model was then validated in the testing set. Variable selection was performed by using the genetic algorithms (GA) [35]. We run 1000 evolution cycles. A maximum of 200 generations were allowed in each one cycle of evolution. The model performance was evaluated by the area under receiver operating characteristic curve (AUC) with 3-fold cross validation. Since each evolution would result in one best fit chromosome, we obtained 1000 chromosomes after GA iteration. A parsimonious representative model was then developed by forward selection procedure to predict the class membership (5-gene class model). Details of the GA are shown in the SDC. The class model was tested in the testing set (n = 137).

Validation in multiple external datasets

The 5-gene class model was used to predict class membership in external datasets. The prognostic value of the 5-gene class model was compared against the sepsis response signature (SRS), Acute Physiology And Chronic Health Evaluation II (APACHE II) and age. The SRS endotype was previously developed using hierarchical clustering by using features with the most variable probes [23]. The presence of SRS1 identifies individuals with an immunosuppressed phenotype that included features of endotoxin tolerance, T-cell exhaustion, and downregulation of human leucocyte antigen (HLA) class II. SRS1 was associated with higher 14-day mortality than was SRS2. APACHE II is a severity-of-disease classification system with a final score of 0 to 71, with higher scores corresponding to more severe disease and a higher risk of death. It is determined within 24 hrs of admission to an intensive care unit (ICU). Because the 5-gene class model was trained on class membership (e.g. using class membership as the response variable), it was not surprising that the 5-gene class model had moderate prognostic performance. In order to explore whether the 5 genes carried prognostic information, we trained a 5-gene mortality model (e.g. using mortality as the response variable) in the dataset GSE65682 and then validated in the non-overlapping external datasets.

Statistics

Predictive value of the 5-gene class model was explored by using logistic regression models. The interactions between predicted class membership and interventions (hydrocortisone versus control; vasopressin versus norepinephrine) were included in the models. To test whether the predictive value was solely explained by severity of illness, we also explored whether there was significant interaction between APACHE II and interventions.

Role of the funding source

The funding sources had no role in the study design, data collection, data analysis, interpretation, or writing of the manuscript.

Results

Datasets

The initial search identified 280 datasets from GEO and ArrayExpress databases. A total of 228 datasets were screened after removal of 52 duplicated datasets. A total of 187 datasets were excluded due to a variety of exclusion criteria including pediatric patients, in vitro experiment, diseases other than sepsis. The remaining 41 datasets were screened by downloading the gene expression and clinical information datasets. After exclusion of 29 datasets due to incomplete expression or outcome data, we finally included 12 datasets for both qualitative and quantitative analysis (Fig. 2). General features of included datasets are shown in Table 1.

k-means clustering analysis

A total of 6392 gene expression values were obtained after merging the 12 datasets. Autoencoder was performed on these features in the dataset GSE65682, resulting in 100 representative features (Figure S1). These autoencoder features were then further reduced to 50 features by univariate filtering on mortality (e.g. variables with $p < 0.05$ for t-test were used for further analysis).

Clustering analysis was performed using the 50 representative features obtained via autoencoder and univariate filtering. The optimal number of clusters was determined to be two by measuring the total within sum of square and average silhouette width (Fig. 3a and b). The two classes can be well separated in the first two major dimensions (Fig. 3c).

There were 321 patients in class 1 (46.9%) and 364 patients in class 2 (53.1%). Patients in Class 1 showed significantly higher mortality rate in the dataset GSE65682 (21.8% [70/321] vs. 12.1% [44/364]; $p < 0.01$ for Chi-square test).

Functional analysis of the two classes

The top 5 most differentially expressed genes ordered by adjusted p value were C14orf159, STOM, MMP8, RPS6KA5 and AKNA (Table S2). Differentially expressed genes used for functional enrichment analysis were calculated as class 1 versus (minus) class 2. Class 1, as compared with class 2, was characterized by immunosuppression that many important pathways were suppressed such as Th1 and Th2 cell differentiation, T cell receptor signaling pathway, DNA-binding transcription factor activity and cell migration (Fig. 4). More visualizations with upset plot (Figure S2), ridge plot (Figure S3) and network plot (Figure S4) are provided in the SDC.

Genetic algorithm to develop a 5-gene class model

The dataset GSE65682 was split into training and testing subsets by 4:1 ratio. Genetic algorithm was applied to the training set to

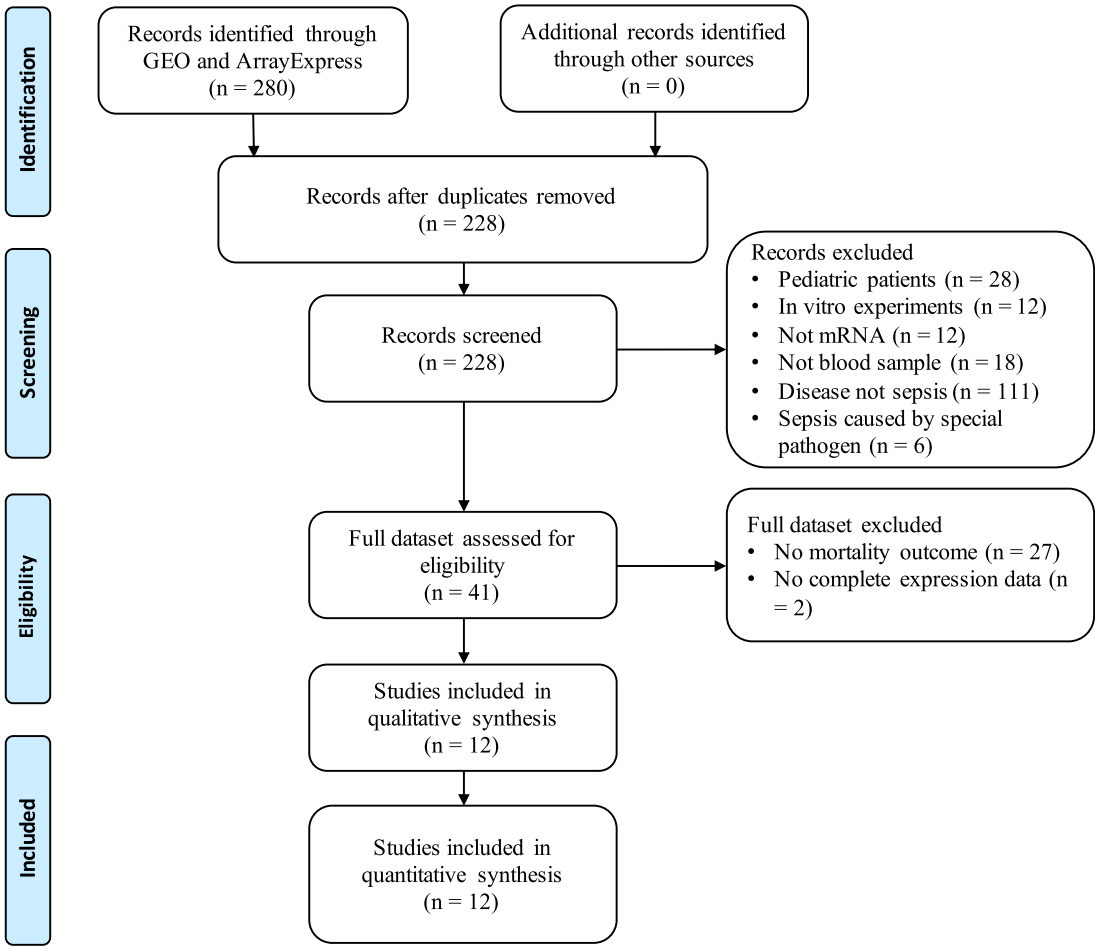

Fig. 2. Flowchart of dataset selection.

Table 1
Dataset included in the study.

Accession	Study population/Inclusion criteria	Cohort description	Timing of gene expression profiling	Country	Timing of mortality	Mortality/Total sample
E-MTAB-7581	Clinical evidence of infection, evidence of a systemic response to infection, and the onset of shock within the previous 72 h (as defined by a systolic blood pressure of <90 mm Hg despite adequate fluid replacement or a need for vasopressors for at least 1 hour) and hypoperfusion or organ dysfunction attributable to sepsis.	Post-hoc of an RCT comparing Vasopressin vs. Norepinephrine for septic shock	At enrollment	United Kingdom	28-day	48/176
E-MTAB-5274	Sepsis defined by ACCP/SCCM guidelines. FP was diagnosed at laparotomy as inflammation of the peritoneal membrane secondary to large bowel perforation and fecal contamination.	Sepsis due to FP	First day of ICU stay	UK	28-day	14/108
E-MTAB-5273	CAP was defined as a febrile illness associated with a cough, sputum production, breathlessness, leukocytosis and radiological features of pneumonia, acquired in the community or within two days of ICU admission.	Sepsis due to CAP, GAinS	First day of ICU stay	UK	28-day	22/118
E-MTAB-4451	The diagnosis of sepsis was based on the International Consensus Criteria with all patients reported here showing some degree of organ dysfunction during ICU admission.	Sepsis due to CAP	On ICU admission	UK	28-day	57/114
GSE65682	ICU patients with suspected CAP for which the attending physician started antibiotic therapy. CAP diagnosis was based on International Sepsis Forum Consensus Conference definition.	Sepsis due to CAP and HAP + non-infectious control	On ICU admission	Netherlands and UK	28-day	114/802
GSE54514	Sepsis was defined as documented bacterial infection in addition to the presence of at least two of the following four clinical criteria: (a) fever or hypothermia (temperature > 38 °C or < 36 °C); (b) tachycardia (>90 beats/min); (c) tachypnea (>20 breaths/min or PaCO2 < 32 mmHg) or the need for mechanical ventilation; (d) an altered white blood cell count of more than 12,000 cells/μL, less than 4000 cells/μL, or the presence of more than 10% band forms	Sepsis	In 24 h of ICU admission	Australia	ICU stay	9/53
GSE63042	Adults at the ED with known or suspected acute infection and presence of at least two SIRS criteria	Sepsis in CAPSOD study	Day of enrollment upon presentation to the ED	United States	28-day	28/106
GSE95233	Septic shock as defined by a systolic blood pressure of <90 mm Hg despite adequate fluid replacement or a need for vasopressors for at least 1 hour.	Septic shock	Day 1 of ICU admission	France	28-day	34/102
GSE106878	Septic shock as defined by a systolic blood pressure of <90 mm Hg despite adequate fluid replacement or a need for vasopressors for at least 1 hour.	Hydrocortisone application in septic shock	Before hydrocortisone application	International	28-day	13/47
GSE33119	Septic shock	activated protein C in septic shock	12 hrs after diagnosis	France	NA	10/20
GSE48080	Patients older than 18 years were enrolled within 48 hrs of the first occurrence of organ dysfunction indicative of severe sepsis or septic shock.	severe sepsis or septic shock	Day 0	Brazil	NA	5/10
GSE33118	Septic shock with a systolic blood pressure of <90 mm Hg despite adequate fluid replacement or a need for vasopressors for at least 1 hour.	NA	Day 0	France	NA	10/20

FP: fecal peritonitis; GAinS: UK Genomic Advances in Sepsis; SIRS: systematic inflammatory response syndrome; CAP: community acquired pneumonia.
Note: Some datasets contained duplicated samples which would be deleted in subsequent analyses.

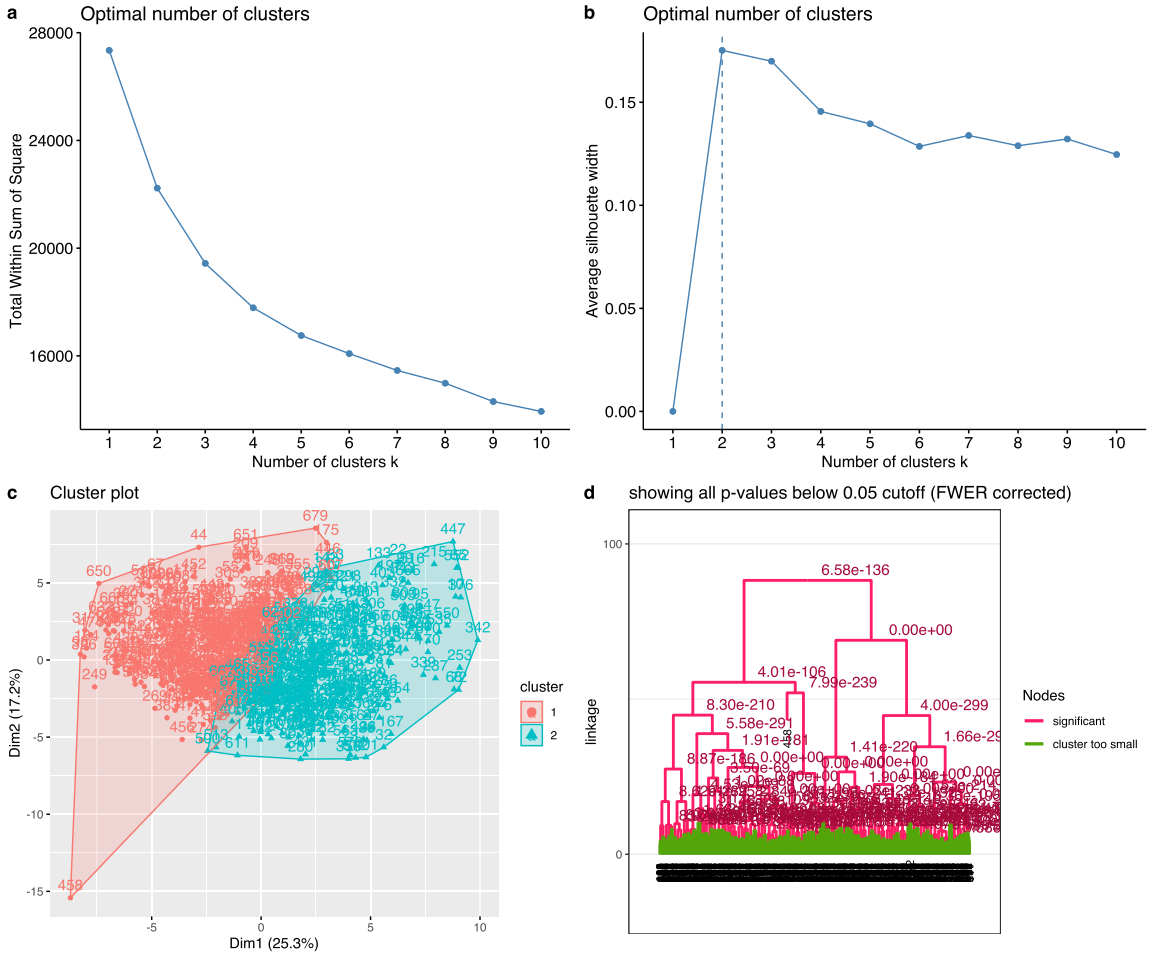

Fig. 3. k-means clustering analysis. a) total within sum of square (WSS) plotted against the number of clusters. Note that the WSS dropped rapidly from 1 to 2 classes. The dropping rate flattened after $k = 2$. b) Average silhouette width plotted against the number of clusters, which indicated that the 2-cluster was the best choice. c) Distribution of subjects in the two most important dimensions. d) Hierarchical clustering also indicated that the two-class model was appropriate. P values for statistical significance were obtained by Monte Carlo based approach. The approach was implemented as a sequential testing procedure guaranteeing control of the family-wise error rate (FWER).

develop a neural network model (5-gene class model). The most frequently selected genes were C14orf159, AKNA, PILRA and STOM (Figure S5). The convergence of the GA was determined by the analysis of the frequency of each gene appeared in the chromosome population. As chromosomes selected by the frequency of each gene in the population will change until no new solutions are found. Therefore, monitoring the stability of gene ranks (based on their frequency) offers the possibility to visualize model convergence. Our results showed that the rank of the top 4 genes (C14orf159, AKNA, PILRA and STOM) were stabilized after 100 evolution cycles, and the fifth ranked gene (USP4) was stabilized after approximately 600 evolution cycles. A representative model comprising 5 genes (C14orf159, AKNA, PILRA, STOM and USP4) was developed by forward selection procedure (Figure S6). When the model was evaluated in the testing set (n = 137), the AUC of the 5-gene class model in predicting class membership was 0.960 (95% CI: 0.932 – 0.988).

Biological functions of the 5 genes

C14orf159 is also known as DGLUCY and is a protein-coding gene. The protein is D-glutamate cyclase that converts D-glutamate to 5-oxo-D-proline [36]. There is also evidence that loss of C14orf159 is associated with the progression of gastric cancer [37]. The product of AKNA may act as a transcription factor that specifically activates the expression of the CD40 receptor and its ligand CD40L/CD154, two cell surface molecules on lymphocytes that are critical for antigen-dependent-B-cell development [38]. PILRA is thought to act as a cellular signaling inhibitory receptor by recruiting cytoplasmic phosphatases like PTPN6/SHP-1 and PTPN11/SHP-2 via their SH2 domains that block signal transduction through dephosphorylation of signaling molecules [39]. STOM regulates ion channel activity and transmembrane ion transport and regulates ASIC2 and ASIC3 channel activity [40]. USP4 (Ubiquitin Specific Peptidase 4) is able to remove conjugated ubiquitin from target proteins [41, 42].

External validation in independent datasets

This section validated our 5-gene class model in the external datasets (n = 780). There were 53 duplicated samples in the datasets E-MTAB-4451 and E-MTAB-5273, which were excluded from analysis. The AUC of the 5-gene class model to predict mortality outcome was 0.707 (95% CI: 0.664 – 0.750), which was better than the age (AUC: 0.572; 95% CI: 0.524 – 0.619). Because the 5-gene class model was not trained on mortality outcome, it is not surprising that the prognostic performance was less than satisfactory. We further trained a model (5-gene mortality model) on mortality with the same 5 genes to examine whether the 5 genes carried prognostic information. The mortality model was trained on the training set (n = 548). When the 5-gene mortality model was tested in the non-overlapping external datasets (n = 780), the AUC was 0.889 (95% CI: 0.861 - 0.917; Fig. 5a). The 5-gene class model performed better than the APACH II score in

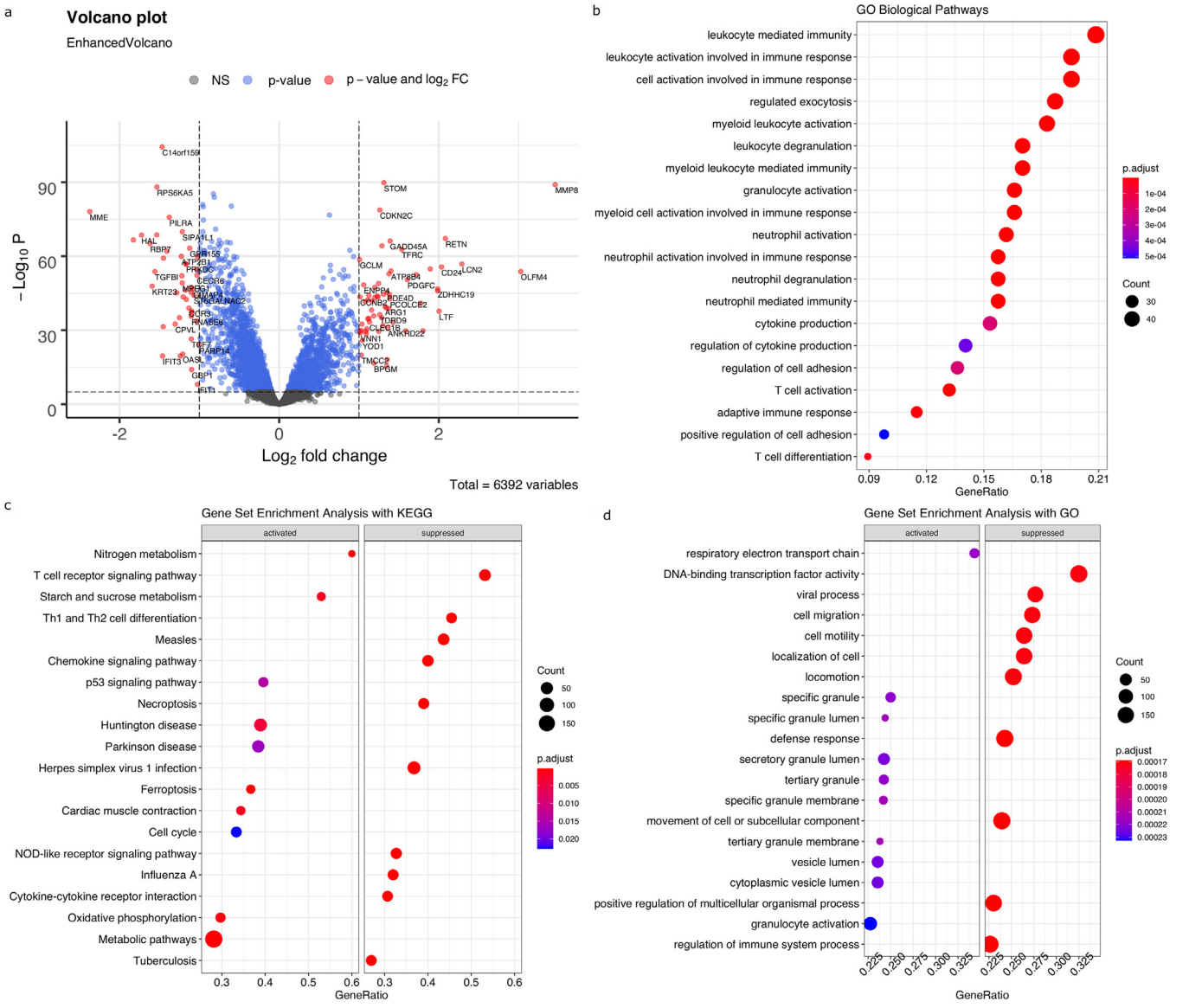

Fig. 4. Functional enrichment analysis for differentially expressed genes between class 1 and 2. a) volcano plot showing differentially expressed genes. Genes with greater than 1 log2 fold changes are annotated. The difference was calculated as Class1- Class2. For example, MMP8 was over expressed in class 1 versus class 2; whereas MME was under-expressed in class 1 versus class 2. b) GO biological pathway enrichment (overrepresentation analysis). c) KEGG gene set enrichment analysis showed that nitrogen metabolism pathway was activated and T cell receptor signaling pathway was suppressed in class 1 versus class 2. d) GO gene set enrichment analysis.

the dataset E-MTAB-7581 (Fig. 5b), the SRS reported in previous studies (Fig. 5c and d).

Predictive value of the 5-gene class model

The dataset E-MTAB-7581 provided information on the use of vasopressin versus norepinephrine, and hydrocortisone versus placebo. The dataset was collected from the VANISH randomized trial with patients randomized to receive either norepinephrine or vasopressin followed by hydrocortisone or placebo [43]. The randomization procedure guarantees the comparability between treated and control groups. We tested whether there was interaction between class membership and the treatment in logistic regression models, by using the binary mortality outcome as the response variable. The result showed that there was no modification effect of both SRS and our Class membership for vasopressin (Table 2). In class 2, the use of hydrocortisone was associated with increased risk of mortality (OR: 3.15 (1.13, 8.82); $p = 0.029$ for likelihood ratio test). The use of hydrocortisone was not associated with changed mortality (OR: 1.88 (0.70, 5.09); $p = 0.211$ for likelihood ratio test) in class 1. SRS overlapped with our classification system in cross tabulation (Table 3).

Discussion

Our study analyzed 12 datasets with whole blood gene expression profiling. We firstly identified two classes of sepsis: class 1 was characterized by immunosuppression and higher mortality rate than class 2. We then developed a 5-gene class model to predict the class membership using genetic algorithms and validated this parsimonious model in external datasets. Our 5-gene class model showed higher prognostic performance than age, APACHE II and SRS. More importantly, the class membership designated by the 5-gene class model had modification effect on hydrocortisone treatment. Our study supports the notion that sepsis is a heterogenous syndrome. Two subclasses with distinct inflammatory responses were identified in the discovery cohort, which is consistent with previous studies [22, 23]. Furthermore, our study used autoencoder to better capture the gene expression features of different endotypes of sepsis. By filtering on

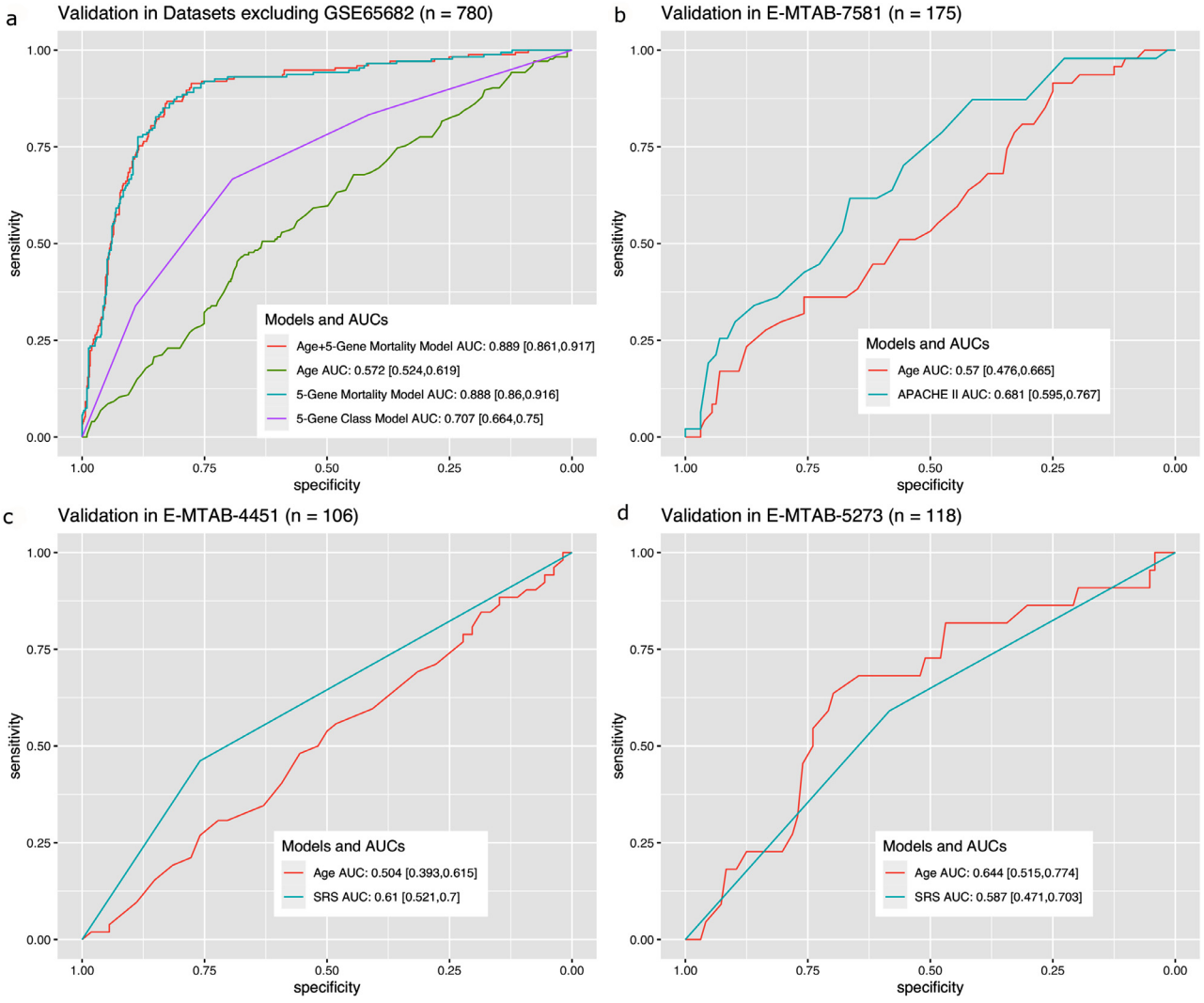

Fig. 5. External validation of the 5-gene class model. Because not all datasets reported the SRS classification, the performance of SRS was reported in individual datasets.
a) the evaluation of the performance of the 5-gene class model in Datasets excluding GSE65682 (n = 780). b) Validation of the 5-gene mortality model in E-MTAB-7581 (n = 175), which showed significantly better discrimination than the APACHE II score. c) The performance of SRS in E-MTAB-4451 (n = 106). d) The performance of SRS in E-MTAB-5273 (n = 118).

mortality before k-means clustering, our model had better prognostic performance than the SRS classification system. Finally, we developed a parsimonious 5-gene class model for the ease of clinical utility. Our preliminary post-hoc analysis showed there was modification effect for the use of hydrocortisone, which was one step closer to the individualized treatment.

Endotypes of sepsis have been widely studied in the literature. The UK Genomic Advances in Sepsis (GAinS) study included critically ill patients with sepsis due to community-acquired pneumonia (CAP) and identified two subclasses SRS1 and SRS2 [23], which showed distinct inflammatory responses to infectious stimuli. Similar SRS endotypes were replicated in sepsis caused by fecal peritonitis [22]. This SRS group membership could not modify the effect of vasopressin versus norepinephrine [43]. Consistent with the GAinS study, our results showed that although the newly developed classification system could not modify the effect of vasopressin versus norepinephrine, class 2 showed significantly higher mortality risk when assigned to the hydrocortisone group. Class 2 is relatively immunocompetent with lower mortality rate as compared to class 1. The use of hydrocortisone suppresses the immune system [44, 45], whereby transforming them to the class 1 group of immunosuppression, and the mortality rate might be increased. This can explain the modification effect of the newly developed classification system for the application of hydrocortisone. Another study utilized cohort from the Molecular Diagnosis and Risk Stratification of Sepsis (MARS) project to perform consensus clustering. They identified four subclasses of sepsis [46]. However, that study only reported that these subclasses had different survival outcomes, further clinical utility of the MARS classification was not investigated. The difference in our study is that we used representative features derived from autoencoder, rather than the original gene expression value, to perform clustering analysis. We showed that the prognostic and predictive value of our 5-gene class model had potential clinical utility.

As compared with previous studies to identify endotypes of sepsis, our study developed a classification system for sepsis with various causes, under the hypothesis that different infection sites can cause systematic inflammatory response, leading to sepsis-related organ dysfunction. The final common pathway is similar across different clinical phenotypes. Thus, we employed the largest sepsis dataset from the MARS consortium (also contained a validation cohort from the GAinS study), containing sepsis from various causes. We proposed that a prediction model developed in this way can have more generalizability to sepsis with various causes. This is supported

Table 2
Comparisons of the Predictive value of 5-gene class model versus disease severity and SRS.

Models	OR (95% CI)	p
Use of hydrocortisone in Class 2	3.15 [1.13, 8.82]	0.029
Use of hydrocortisone in Class 1	1.88 [0.70, 5.09]	0.211
Use of hydrocortisone in SRS 2	3.76 [1.41, 10.04]	0.008
Use of hydrocortisone in SRS 1	1.25 [0.47, 3.36]	0.658
Use of hydrocortisone by APACHE II	0.94 [0.86, 1.03]	0.210
Use of vasopressin in SRS 2	0.69 [0.18, 2.62]	0.583
Use of vasopressin in SRS 1	1.50 [0.40, 3.89]	0.403
Use of vasopressin in Class 2	2.52 [0.62, 10.74]	0.201
Use of vasopressin in Class 1	1.51 [0.54, 4.24]	0.433
Use of vasopressin by APACHE II	0.96 [0.87, 1.06]	0.427

The logistic regression models integrating interactions between treatment allocation and SRS, Class or APACHE II were built in the dataset E-MTAB-7581. A total of 6 logistic regression models were built by using mortality as the response variable and respective predictors and interactions were: hydrocortisone × *Class*, hydrocortisone × *SRS*, hydrocortisone × *APACHE II*, vasopressin × *Class*, vasopressin × *SRS*, vasopressin × *APACHE II*. A significant ($p < 0.05$) interaction indicated that the classification method was of predictive value because it identifies a subgroup of patients respond differently to treatment.
SRS classification was used as previously reported.
The 5-gene class model developed and validated in GSE65682 was used to classify patients into Class 1 and Class 2.
Abbreviations: SRS: sepsis response signature; APACHE II: Acute Physiology and Chronic Health Evaluation.

by the finding that the 5-gene class model performed better than SRS in predicting mortality in external independent cohorts.

Our study has several limitations. First, many publicly available sepsis datasets did not report mortality outcome; these studies primarily focused on the differential diagnosis of sepsis versus non-infectious systematic inflammatory response syndrome. The exclusion of these datasets may introduce potential selection bias. Second, many genes were excluded during the merging of datasets, resulting in the loss of some important genes. However, because we needed to validate our model in independent datasets, we must ensure that the genes used for model construction were also available in the testing datasets. Third, the dataset used for investigation of the interaction between class membership and treatment was of limited sample size. Large sample size can increase the statistical power to detect clinically meaningful modification effect by subclasses. Forth, there is a possibility that the sepsis is not categorized by the 5 genes. Rather, the categorization might reflect the progression (stage) of the disease. In other words, the results may not be discoveries of different kinds of sepsis but may be observation of different stages of the disease. In the case of mortality prediction, it might mean that the treatment was too late. Finally, the sepsis patients included in our analysis were not guaranteed to be free of other diseases. However, since the original datasets did not provide full details of other diseases/comorbidities, the impact of other diseases on the predictive performance of our model cannot be fully addressed. The study is also limited by the relative sparsity of individual instances versus the potential feature space. We have tried to do the primary testing on the more parsimonious set, but the problem of overfitting is still present.

Table 3
Cross tabulation of subclass membership for Class and SRS classification system.

	Class 1	Class 2
SRS 1	64	29
SRS 2	34	49

Abbreviations: SRS: sepsis response signature.

In conclusion, our study identified two classes of sepsis that showed different mortality outcome and response to hydrocortisone therapy. Class 1 was characterized by immunosuppression with higher mortality rate; whereas class 2 was relatively immunocompetent. We further developed a 5-gene class model to predict class membership.

Contributors

ZZ conceived the idea and drafted the manuscript; QP interpreted the results and deep learning method; HG and LX performed data collection and curation; YH and PC performed some interpretation of the results and. All authors read and approved the final version of the manuscript.

Declaration of Competing Interest

There is no conflict of interest.

Acknowledgements

The study was funded by the National Natural Science Foundation of China (Grant No. 81901929).

Data Sharing Statement

All data were available in publicly accessible databases. Accession numbers were available in Table 1.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ebiom.2020.103081.

Reference

[1] László I, Trásy D, Molnar Z, Fazakas J. Sepsis: from Pathophysiology to Individualized Patient Care. J Immunol Res 2015;2015 510436–13. doi: 10.1155/2015/510436.
[2] Leligdowicz A, Matthay MA. Heterogeneity in sepsis: new biological evidence with clinical applications. Crit Care 2019;23 80–8. doi: 10.1186/s13054-019-2372-2.
[3] Seymour CW, Kennedy JN, Wang S, Chang C-CH, Elliott CF, Xu Z, Berry S, Clermont G, Cooper G, Gomez H, Huang DT, Kellum JA, Mi Q, Opal SM, Talisa V, van der Poll T, Visweswaran S, Vodovotz Y, Weiss JC, Yealy DM, Yende S, Angus DC. Derivation, Validation, and Potential Treatment Implications of Novel Clinical Phenotypes for Sepsis. JAMA 2019;321:2003–17. doi: 10.1001/jama.2019.5791.
[4] Andrews B, Semler MW, Muchemwa L, Kelly P, Lakhi S, Heimburger DC, Mabula C, Bwalya M, Bernard GR. Effect of an Early Resuscitation Protocol on In-hospital Mortality Among Adults With Sepsis and Hypotension: a Randomized Clinical Trial. JAMA 2017;318:1233–40. doi: 10.1001/jama.2017.10913.
[5] Kawazoe Y, Miyamoto K, Morimoto T, Yamamoto T, Fuke A, Hashimoto A, Koami H, Beppu S, Katayama Y, Itoh M, Ohta Y, Yamamura H, Dexmedetomidine for Sepsis in Intensive Care Unit Randomized Evaluation (DESIRE) Trial Investigators. Effect of Dexmedetomidine on Mortality and Ventilator-Free Days in Patients Requiring Mechanical Ventilation With Sepsis: a Randomized Clinical Trial. JAMA 2017;317:1321–8. doi: 10.1001/jama.2017.2088.
[6] National Heart, Lung, and Blood Institute ARDS Clinical Trials Network, Truwit JD, Bernard GR, Steingrub J, Matthay MA, Liu KD, Albertson TE, Brower RG, Shanholtz C, Rock P, Douglas IS, DeBoisblanc BP, Hough CL, Hite RD, Thompson BT. Rosuvastatin for sepsis-associated acute respiratory distress syndrome. N. Engl. J. Med. 2014;370:2191–200. doi: 10.1056/NEJMoa1401520.
[7] Seymour CW, Gesten F, Prescott HC, Friedrich ME, Iwashyna TJ, Phillips GS, Lemeshow S, Osborn T, Terry KM, Levy MM. Time to Treatment and Mortality during Mandated Emergency Care for Sepsis. N. Engl. J. Med. 2017;376:2235–44. doi: 10.1056/NEJMoa1703058.
[8] Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. Nat. Med. 2018;24:1716–20. doi: 10.1038/s41591-018-0213-5.
[9] Peng X, Ding Y, Wihl D, Gottesman O, Komorowski M, Lehman L-WH, Ross A, Faisal A, Doshi-Velez F. Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning. AMIA Annu Symp Proc 2018;2018:887–96.
[10] Zhang Z, Zhang G, Goyal H, Mo L, Hong Y. Identification of subclasses of sepsis that showed different clinical outcomes and responses to amount of fluid

resuscitation: a latent profile analysis. Crit Care 2018;22:347. doi: 10.1186/s13054-018-2279-3.
[11] Bhavani SV, Carey KA, Gilbert ER, Afshar M, Verhoef PA, Churpek MM. Identifying Novel Sepsis Subphenotypes Using Temperature Trajectories. Am. J. Respir. Crit. Care Med 2019;200:327–35. doi: 10.1164/rccm.201806-1197OC.
[12] Wiersema R, Jukarainen S, Vaara ST, Poukkanen M, Lakkisto P, Wong H, Linder A, van der Horst ICC, Pettilä V. Two subphenotypes of septic acute kidney injury are associated with different 90-day mortality and renal recovery. Crit Care 2020;24 150–10. doi: 10.1186/s13054-020-02866-x.
[13] Zhai J, Qi A, Zhang Y, Jiao L, Liu Y, Shou S. Bioinformatics Analysis for Multiple Gene Expression Profiles in Sepsis. Med. Sci. Monit. 2020;26:e920818. doi: 10.12659/MSM.920818.
[14] Ballman KV. Biomarker: predictive or Prognostic? J. Clin. Oncol 2015;33:3968–71. doi: 10.1200/JCO.2015.63.3651.
[15] Sechidis K, Papangelou K, Metcalfe PD, Svensson D, Weatherall J, Brown G. Distinguishing prognostic and predictive biomarkers: an information theoretic approach. Bioinformatics 2018;34:3365–76. doi: 10.1093/bioinformatics/bty357.
[16] Xu D, Liao S, Li P, Zhang Q, Lv Y, Fu X, Yang M, Wang J, Kong L. Metabolomics Coupled with Transcriptomics Approach Deciphering Age Relevance in Sepsis. Aging Dis 2019;10:854–70. doi: 10.14336/AD.2018.1027.
[17] Zador Z, Landry A, Balas M, Marshall JC, Cusimano MD. Data Driven Analysis Reveals Shared Transcriptome Response, Immune Cell Composition, and Distinct Mortality Rates Across Differing Etiologies of Critical Illness. Crit. Care Med. 2020;48:338–43. doi: 10.1097/CCM.0000000000004128.
[18] 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 287ra71–287ra71. doi: 10.1126/scitranslmed.aaa5993.
[19] Cheng L, Nan C, Kang L, Zhang N, Liu S, Chen H, Hong C, Chen Y, Liang Z, Liu X. Whole blood transcriptomic investigation identifies long non-coding RNAs as regulators in sepsis. J Transl Med 2020;18 217–13. doi: 10.1186/s12967-020-02372-2.
[20] Schaack D, Siegler BH, Tamulyte S, Weigand MA, Uhle F. The immunosuppressive face of sepsis early on intensive care unit-A large-scale microarray meta-analysis. PLoS ONE 2018;13:e0198555. doi: 10.1371/journal.pone.0198555.
[21] Mayhew MB, Buturovic L, Luethy R, Midic U, Moore AR, Roque JA, Shaller BD, Asuni T, Rawling D, Remmel M, Choi K, Wacker J, Khatri P, Rogers AJ, Sweeney TE. A generalizable 29-mRNA neural-network classifier for acute bacterial and viral infections. Nat Commun 2020;11 1177–10. doi: 10.1038/s41467-020-14975-w.
[22] Burnham KL, Davenport EE, Radhakrishnan J, Humburg P, Gordon AC, Hutton P, Svoren-Jabalera E, Garrard C, Hill AVS, Hinds CJ, Knight JC. Shared and Distinct Aspects of the Sepsis Transcriptomic Response to Fecal Peritonitis and Pneumonia. Am. J. Respir. Crit. Care Med. 2017;196:328–39. doi: 10.1164/rccm.201608-1685OC.
[23] Davenport EE, Burnham KL, Radhakrishnan J, Humburg P, Hutton P, Mills TC, Rautanen A, Gordon AC, Garrard C, Hill AVS, Hinds CJ, Knight JC. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. Lancet Respir Med 2016;4:259–71. doi: 10.1016/S2213-2600(16)00046-1.
[24] Sweeney TE, Azad TD, Donato M, Haynes WA, Perumal TM, Henao R, Bermejo-Martin JF, Almansa R, Tamayo E, Howrylak JA, Choi A, Parnell GP, Tang B, Nichols M, Woods CW, Ginsburg GS, Kingsmore SF, Omberg L, Mangravite LM, Wong HR, Tsalik EL, Langley RJ, Khatri P. Unsupervised Analysis of Transcriptomics in Bacterial Sepsis Across Multiple Datasets Reveals Three Robust Clusters. Crit. Care Med. Publish Ahead of Print 2018;1. doi: 10.1097/CCM.0000000000003084.
[25] Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007;8:118–27. doi: 10.1093/biostatistics/kxj037.
[26] Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28:882–3. doi: 10.1093/bioinformatics/bts034.
[27] Bank D, Koenigstein N, Giryes R. Autoencoders. arXiv 2020.
[28] Huang Z, Johnson TS, Han Z, Helm B, Cao S, Zhang C, Salama P, Rizkalla M, Yu CY, Cheng J, Xiang S, Zhan X, Zhang J, Huang K. Deep learning-based cancer survival prognosis from RNA-seq data: approaches and evaluations. BMC Med Genomics 2020;13:41. doi: 10.1186/s12920-020-0686-1.
[29] Zhang S, Li X, Lin Q, Lin J, Wong K-C. Uncovering the key dimensions of high-throughput biomolecular data using deep learning. Nucleic Acids Res 2020;48 e56–e56. doi: 10.1093/nar/gkaa191.
[30] Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: an RPackage for Determining the Relevant Number of Clusters in a Data Set. J Stat Softw 2014;61. doi: 10.18637/jss.v061.i06.
[31] Kimes PK, Liu Y, Hayes DN, Marron JS. Statistical Significance for Hierarchical Clustering. arXiv. 2014.
[32] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43 e47–e47. doi: 10.1093/nar/gkv007.
[33] Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16:284–7. doi: 10.1089/omi.2011.0118.
[34] Sinha P, Delucchi KL, McAuley DF, O'Kane CM, Matthay MA, Calfee CS. Development and validation of parsimonious algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. Lancet Respir Med 2020;8:247–57. doi: 10.1016/S2213-2600(19)30369-8.
[35] Trevino V, Falciani F. GALGO: an R package for multivariate variable selection using genetic algorithms. Bioinformatics 2006;22:1154–6. doi: 10.1093/bioinformatics/btl074.
[36] Ariyoshi M, Katane M, Hamase K, Miyoshi Y, Nakane M, Hoshino A, Okawa Y, Mita Y, Kaimoto S, Uchihashi M, Fukai K, Ono K, Tateishi S, Hato D, Yamanaka R, Honda S, Fushimura Y, Iwai-Kanai E, Ishihara N, Mita M, Homma H, Matoba S. Glutamate is metabolized in the heart mitochondria. Sci Rep 2017;7 43911–9. doi: 10.1038/srep43911.
[37] Zhu Y-M, Li Q, Gao X-Z, Meng X, Sun L-L, Shi Y, Lu E-T, Zhang Y. C14orf159 suppresses gastric cancer cells' invasion and proliferation by inactivating ERK signaling. Cancer Manag Res 2019;11:1717–23. doi: 10.2147/CMAR.S176771.
[38] Siddiqa A, Sims-Mourtada JC, Guzman-Rojas L, Rangel R, Guret C, Madrid-Marina V, Sun Y, Martinez-Valdez H. Regulation of CD40 and CD40 ligand by the AT-hook transcription factor AKNA. Nature 2001;410:383–7. doi: 10.1038/35066602.
[39] Mousseau DD, Banville D, L'Abbé D, Bouchard P, Shen SH. PILRalpha, a novel immunoreceptor tyrosine-based inhibitory motif-bearing protein, recruits SHP-1 upon tyrosine phosphorylation and is paired with the truncated counterpart PILRbeta. J. Biol. Chem. 2000;275:4467–74. doi: 10.1074/jbc.275.6.4467.
[40] Mairhofer M, Steiner M, Salzer U, Prohaska R. Stomatin-like protein-1 interacts with stomatin and is targeted to late endosomes. J. Biol. Chem. 2009;284:29218–29. doi: 10.1074/jbc.M109.014993.
[41] Clerici M, Luna-Vargas MPA, Faesen AC, Sixma TK. The DUSP-Ubl domain of USP4 enhances its catalytic efficiency by promoting ubiquitin exchange. Nat Commun 2014;5 5399–11. doi: 10.1038/ncomms6399.
[42] Wada K, Tanji K, Kamitani T. Oncogenic protein UnpEL/Usp4 deubiquitinates Ro52 by its isopeptidase activity. Biochem. Biophys. Res. Commun. 2006;339:731–6. doi: 10.1016/j.bbrc.2005.11.076.
[43] Antcliffe DB, Burnham KL, Al-Beidh F, Santhakumaran S, Brett SJ, Hinds CJ, Ashby D, Knight JC, Gordon AC. Transcriptomic Signatures in Sepsis and a Differential Response to Steroids. From the VANISH Randomized Trial. Am. J. Respir. Crit. Care Med. 2019;199:980–6. doi: 10.1164/rccm.201807-1419OC.
[44] Coutinho AE, Chapman KE. The anti-inflammatory and immunosuppressive effects of glucocorticoids, recent developments and mechanistic insights. Mol. Cell. Endocrinol. 2011;335:2–13. doi: 10.1016/j.mce.2010.04.005.
[45] Fauci AS. Mechanisms of the immunosuppressive and anti-inflammatory effects of glucocorticosteroids. J Immunopharmacol 1978;1:1–25. doi: 10.3109/08923977809027327.
[46] Scicluna BP, van Vught LA, Zwinderman AH, Wiewel MA, Davenport EE, Burnham KL, Nürnberg P, Schultz MJ, Horn J, Cremer OL, Bonten MJ, Hinds CJ, Wong HR, Knight JC, van der Poll T, MARS consortium. Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study. Lancet Respir Med 2017;5:816–26. doi: 10.1016/S2213-2600(17)30294-1.