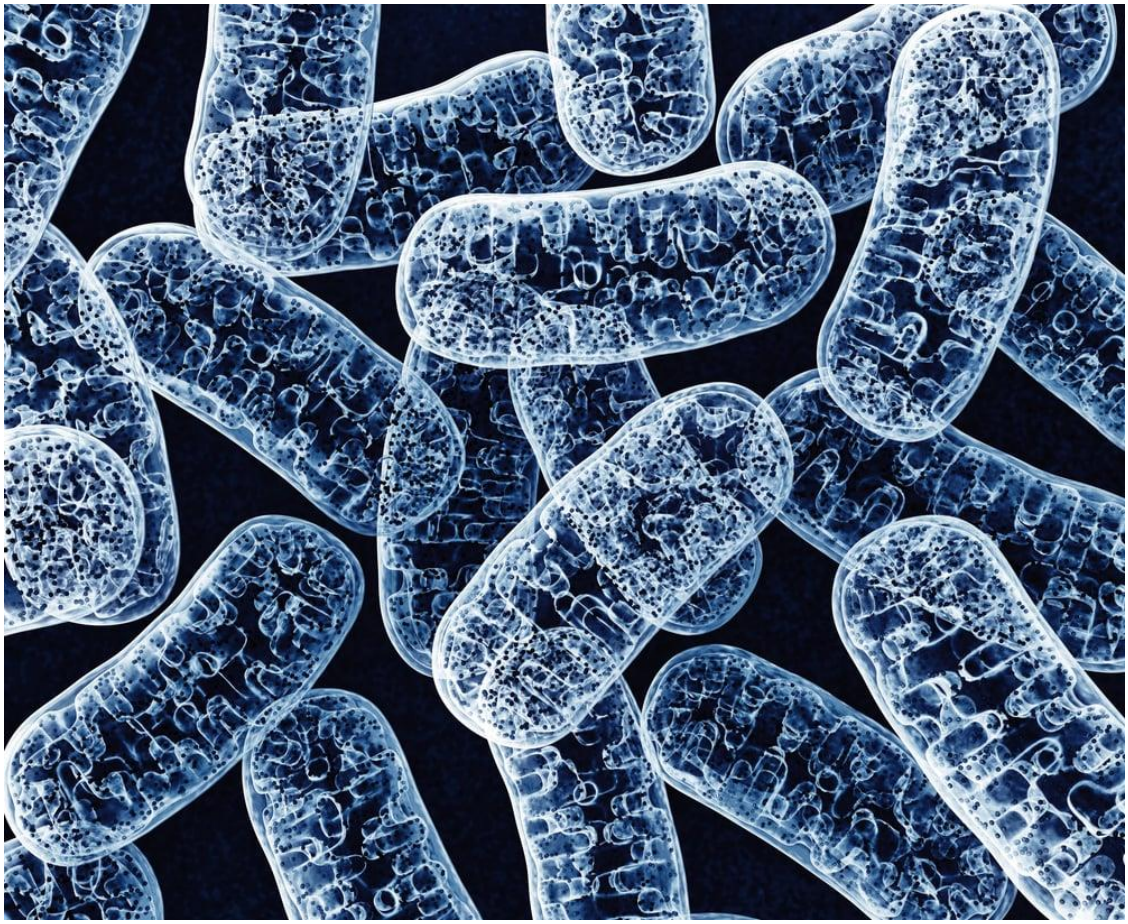
Unraveling the role of mitochondria dysfunction in early sepsis

Identifying and predicting mitochondrial-driven endotypes in early-stage sepsis

Jamie Scheper

03-01-2024



Name: Jamie Scheper
Student ID: 373689
Class: BFV-4
Date: 03-01-2024
Supervisor: Fenna Feenstra

Unraveling the role of mitochondria dysfunction in early sepsis

Identifying and predicting mitochondrial-driven endotypes in early-stage sepsis

Name:	Jamie Scheper
Student ID:	373689
Study:	Bioinformatics (BFV-4)
School:	Hanze UAS
Institute:	Life Sciences & Technology (ILST)
Supervisor Hanze:	Fenna Feenstra
Intern supervisor:	Jingyi Lu and Prof. Dr. H.R. Bouma
Date of completion:	03-01-2024
Status:	Not Confidential

Abstract

Sepsis, a major contributor to in-hospital deaths, is challenging to detect due to its heterogeneous nature. Dysfunctional mitochondria from immune responses can worsen the condition. Analyzing mitochondria-related genes could help in early detection and guide treatment.

RNA-Seq and clinical data from 348 septic patients, originating from four emergency rooms (ER) and one intensive care unit (ICU), were used and compared to 44 healthy controls. Both supervised and unsupervised machine-learning algorithms were applied to identify gene profiles with clinical relevance.

Mitochondria-related differentially expressed genes (DEGs) were identified by comparing levels of severity based on SOFA scores. These DEGs helped establish two unique severity-based on ER cohorts ('endotypes'), distinct from healthy controls and validated on ICU cohort. This categorization helps to address sepsis heterogeneity. Feature selection identified three novel gene sets that could accurately predict endotype group and severity in both ER and ICU settings. Notably, a novel eleven-gene set predicted endotypes with 95% accuracy and severity with 76% accuracy.

The gene profiles and endotypes indicate the essential role mitochondrial genes play in early sepsis detection and as potential novel biomarkers. Future research should aim to develop a multi-modal prognostic tool for enhanced patient stratification.

Abbreviations

Full term	Abbreviation
Absolute neutrophil count	ANS
Adenosine triphosphate	ATP
Area under the curve	AUC
Bisphenol A	BPA
Damage-associated molecular pattern	DAMP
Differentially expressed gene	DEG
Differentially expression	DE
Emergency room	ER
Exploratory data analysis	EDA
Fold change	FC
Intensive care unit	ICU
Kif2-binding protein	KBP
Least absolute shrinkage and selection operator	LASSO
Institute for Life Science and Technology	ILST
Median absolute deviation	MAD
Partition around medoids	PAM
Principal component analysis	PCA
Quick SOFA	qSOFA
Reactive oxygen species	ROS
Receiver operating characteristic	ROC
Recursive feature elimination with cross-validation	RFECV
Sequential feature selection	SFS
Sequential organ failure assessment	SOFA
Variance stabilizing transformation	VST
White blood cell count	WBC
Within-cluster sum of squares	WSS

Organization

This project was provided by Acutelines, a biobank and research group that focuses on acute diseases such as sepsis. The mission of Acutelines is to support acute and emergency medicine research by offering a foundation for novel clinical studies and to enable the development of personalized medicine through data, images, and biomaterials.

This project was supervised by Prof. Dr. H.R. Bouma and daily supervisor Ph.D. Candidate Jingyi Lu.

Table of Content

Introduction

Sepsis is defined as life-threatening organ failure due to a dysregulated host response to infection. [1] Over the past few decades, mortality related to severe sepsis has gradually declined, which can be attributed to the timely use of antibiotics, prevention, and understanding of molecular pathways. [2] [3] However, sepsis remains one of the largest contributors to in-hospital mortality rates and high healthcare costs attributed to the absence of targeted medication despite medical breakthroughs. [4] Patients with sepsis have a high rate of hospital admission; the average length of stay was 75% greater than in other conditions. [5] [6] There were a total of 48.9 million sepsis cases worldwide in which one in five patients perished, which adds up to almost 20% of global deaths. [4]

The variability in disease expression, site of infection, and host response makes early recognition and adequate treatment difficult. Such treatments are currently limited to antibiotics and supportive care. The heterogeneous nature of sepsis was a reason to redefine sepsis as a syndrome, which gives rise to subclasses of sepsis ("endotypes"). Distinct functional mechanisms define endotypes. Previous research indicated the existence of two to five endotypes of sepsis. [7] [8] [9] One such study, conducted by Baquir and coresearchers, attempted to use transcriptional and clinical data to identify gene expression profiles and to stratify patients into five distinct endotypes through unsupervised machine learning techniques. [9] An unique gene profile characterized each endotype and represents a different severity status of sepsis based on sequential organ failure assessment (SOFA) scores. A SOFA score is based on the assessment of six different organ systems, giving each a score from 0 to 4, with higher scores indicating severe dysfunction. A total score of 24 is therefore the maximum score. [10] SOFA scores are an important treatment and diagnostic tool for septic patients. In addition, a quick SOFA (qSOFA) is a simplified version based on three criteria and is used in the ER to give a quick assessment of a patient's health. [11] A qSOFA score of two or higher indicates sepsis and a need for further testing. [11] According to the authors, stratifying patients through whole-blood RNA-Seq may optimize antibiotic treatment and address the heterogeneity of sepsis by identifying predictive biomarkers. This way, more personalized therapeutics can be developed.

Mitochondria play an essential role in cellular homeostasis: producing adenosine triphosphate (ATP) for a continuous energy supply, managing reactive oxygen species (ROS), and regulating inflammatory and apoptosis pathways. [12] During sepsis, an increase in ROS potentially leads to a disturbance in the redox balance of mitochondria. The loss of balance gives rise to a state of oxidative stress, which damages mitochondrial membrane permeability and homeostasis, which can lead to cell death. [13] As a result of apoptosis, mitochondrial damage-associated molecular patterns (DAMPs) are released and contribute to an even larger systemic inflammatory response. [14] There exists a well-defined relationship between multi-organ failure, elevated levels of oxidative stress, and mortality related to sepsis. [15] Recent studies which tried to establish endotypes have focused on immune-related gene expression profiles and proved valuable in paving the way for immune-specific therapeutics. [9] However, mitochondrial dysfunction significantly contributes to the heterogeneity of sepsis and, if not utilized, would represent a substantial oversight. The stratification into specific mitochondrial-driven endotypes may identify

distinct subgroups or even stable biomarkers. In addition, it served as a robust - combined with immune-related endotypes - a predictive tool to enhance early detection of sepsis, open doors to new personalized therapeutics, and ultimately, improve patient outcomes. For example, a recent study regarding the preservation of mitochondrial membrane potential based on early detection of sepsis based on endotypes demonstrated improved outcomes. [16]

This study attempted to associate mitochondrial dysfunction with clinical parameters by identifying mitochondrial-driven endotypes based on analyzing whole-blood transcriptomics. Emergency room (ER) and intensive care unit (ICU) cohorts from the publicly available dataset (GSEA number: GSE185263) were used to identify mitochondrial-driven severity and mortality gene signatures. In addition, clinical parameter were obtained upon request. This study aimed to establish gene signatures by identifying differentially expressed genes (DEGs) by comparing patients' conditions. DEGs form a genetic understanding of how sepsis severity and outcomes manifest in different endotypes. A key objective of this study was the development of a machine-learning prediction tool that could accurately predict mitochondria-driven endotypes and the severity of a patient's condition based on a refined set of gene profiles.

Materials and Methods

Study population

Recruitment of patients leaned mainly on the suspicion of sepsis based on at least two sepsis-1 criteria and the professional opinion of medical doctors. Patients were at least 18 years of age.

The number of patients with suspected sepsis added up to 348 after preprocessing and was derived from a multi-center study: 266 patients originated from four different ERs (Vancouver, Canada, Colombia, the Netherlands, and Australia), and 82 patients came from one ICU (Toronto, Canada). Additionally, 44 healthy control samples were recruited from three different hospital biobanks (Sydney, Neiva, and Toronto).

Essential clinical variables were sepsis severity and mortality. The sepsis severity variable consists of three levels: 'High' (≥ 5 SOFA), 'Intermediate' (≥2 & <5 SOFA), and 'Low' (< 2 SOFA). Mortality data included survival, fatalities, and cases with unknown outcomes. Clinical variables were amassed to a total of 268, with 116 being ER-only and 69 unique to the ICU cohort. Data contains multiple variables regarding SOFA scores, including for ER, qSOFA, and the worst within 72 hours after admission SOFA and ICU used, in addition to SOFA and APACHE scores. However, the worst within 72 hours SOFA score was used as a benchmark for the sepsis severity variable. The interest lies in clinical variables that indicate mitochondria dysfunction, such as lactate levels, neutrophil counts, oxygen treatment, length of admission, and microbiological results (see Table 1).

Parameter	**ER (*n* = 266)**	**ICU (*n* = 82)**
Age	56,06 ± 1,26 (99,63%)	61,41 ± 1,70 (98,78%)
Gender	F: 45,49%; M: 54,51%	F: 30,49%; M: 68,29%
In-hospital mortality*	D: 12,03%; S: 87,59%	Not available
Severity**	H: 12,03%; I: 40,98%; L: 47,00%	H: 63,41%; I: 20,73%; L: 15,85%
ER qSOFA	1,97 ± 0,12 (100%)	Not relevant
SOFA	1,31 ± 0,14 (100%)	7,34 ± 0,55 (100%)
Lactate	2,24 ± 0,25 (18,05%)	Not available
Neutrophil count	6,59 ± 0,41 (56,02%)	Not available
Total white blood cell count	9,39 ± 0,39 (80,45%)	9,48 ± 0,67 (98,78%)
Total admission (in days)	7,53 ± 0,54 (97,74%)	12,04 ± 0,97 (100%)

*Levels mortality: Deceased (D) and Survived (S).
**Levels severity: High (H), Intermediate (I), and Low (L).

Table 1—Essential clinical parameters related to mitochondria dysfunction and sepsis severity of ER and ICU cohorts. Here, the mean and (±) standard error of numeric variables are depicted. For categorical variables, the distribution per level is presented. For both numeric and categorical variables, the number of total records that are available is presented in percentage (%).

RNA data collection

2.5 ml of blood for RNA sequencing was collected within 72 hours after admission, and RNA-Seq was performed with Illumina NovaSeq 6000 S4, which resulted in 100 long base pair reads. Thereafter, 'fastqc' (version 0.11.9) and 'multiqc' (version 1.6) were used for quality control, alignment was conducted with Ensembl's 'STAR' (version 2.7.9a) and read count assessment with 'htseq' (version 0.11.3). [17] [18] [19] [20] RNA-Seq data was composed of roughly 58.000 genes, with 1.690 being related to mitochondria, determined by the definition used in the MSigDB database (via ID "GO:0005739").

Exploratory data analysis

The clinical and count datasets were analyzed in R (version 4.3.1). [21] The focus was on cleaning the data from missing values, either by removal or imputation, taking out attributes that would not contribute to our research goals, adjusting the structure of valuable attributes, highlighting the differences between cohorts, and getting to know the data structure. We used many external packages to explore the data. For example, the packages 'dplyr' (version 1.1.3) and 'tidyverse' (version 2.0.0) were used extensively to mutate data, remove redundancies, and explore the correlations between attributes. [22] [23] Other external packages had a minor function; for example, 'caret' (version 6.0-94) was used to remove near-zero variance attributes. [24] For visualization purposes, 'ggplot2' (version 3.4.3) was used throughout the project. [25] All transformations regarding the data structure are thoroughly explained, and choices are discussed in the supplied Exploratory Data Analysis (EDA), which can be found, as all research logs, in the accommodating BitBucket repository. [26]

Differential expression and pathway analysis

Identification of differentially expressed genes was undertaken with 'DESeq2' (version 1.42.0) and, to validate our findings, 'egdeR' (version 4.0.2). [27] [28] We compared gene expression patients from both the ER and ICU based on sepsis severity status and mortality outcome. Low-expressed genes were filtered out if counts were less than ten over at least ten samples. In addition, genes were considered a DEG when displaying an adjusted p-value of ≤ 0.05 based on a Wald test and ≥ 1.2 absolute fold change (FC). Due to this low amount of DEGs in mortality, we primarily focused on establishing severity signatures. We compared each level from the sepsis severity variable to each other and extracted each unique DEG across all comparisons.

Pathways in which DEGs reside were identified by using the ReactomePA (version 1.46.0 with database 'Reactome') and MSigDB Hallmark database via the R-based 'enrichR' (version 3.2) package. [29] [30] Pathways were considered significant when they exhibited an adjusted p-value of ≤ 0.05 (a hypergeometric test via enrichR's internal operations).

DESeq2's internal normalization efforts account for variations in, for example, sequencing depth. To ensure a similar correction was applied for downstream analysis, we transformed the raw count data using DESeq2's variance stabilizing transformation (VST) with a batch correction using the function 'ComBat' from R package 'SVA' (version 3.50.0). [31]

Unsupervised clustering

Unsupervised clustering machine learning models were employed to establish severity-based endotypes. We looked at three different methods, namely K-medoids (partition around medoids (PAM)), K-means, and hierarchical clustering. We used to build dendrograms for hierarchical clustering via 'dendextend' (version 1.17.1). [32] For each unsupervised clustering method, we compared between different distances and, where appropriate, linkage methods. Determining cluster stability through several metrics such as average silhouette score, elbow plots, and consensus clustering via the R-based package 'ConsensusClusterPlus' (version 1.66.0). [33] We determined the optimal amount of clusters based on the metric results of within-cluster sum of squares (WSS), silhouette scores, and GAP scores via packages 'factoextra' (version 2.9) and build-in 'cluster' (version 2.1.6); and 30 different cluster scoring indices of the R package 'NbClust' (version 3.0.1) that give the optimal amount of clusters through majority vote. [34] [35] Thereafter, we compared these endotypes by performing various DE and pathway analyses between endotype groups, amongst one another in a one-versus-rest approach and comparing them with healthy control samples while using the same parameter settings as described above. Comparisons of clinical variables across endotypes were conducted using Wilcoxon rank-sum tests for numeric attributes and Chi-squared tests for categorical variables, followed by a Bonferroni correction for post-hoc analysis, respectively. Heatmap visualizations were done with the package 'ComplexHeatmap' (version 2.18.0). [36]

Gene set and prediction model

For this section, we used the Python (version 3.11.2) language and Jupyter for its notebook integration (version 1.0.0). [37] [38] Data manipulation was done with 'pandas' (version 2.1.3) and NumPy (version 1.26.2). [39] [40] After establishing severity-related endotypes, we considered four different supervised machine learning models to predict the cluster number and severity status of patients. Classification based on support vector machine (SVM), random forest, naïve Bayes, and logistic regression (with L1 or least absolute shrinkage and selection operator (LASSO) regularization) was undertaken for the entire DEG set. We accomplished this by using the machine-learning package 'sklearn' (version 1.3.2; also known as 'scikit-learn'). [41] We measured the performance of the included classification models with scoring metrics such as F1, accuracy, area under the curve (AUC), and recall scores. In addition, we plotted a receiver operating characteristic (ROC) curve for every comparison in a one-versus-rest fashion when dealing with more than two groups. In addition, 'seaborn', (version 0.13.0) and 'matplotlib' (version 3.8.2), 'scikit-plot' (version

0.3.7) were used for visualization purposes. [42] [43] [44] We trained models on 80% of ER patients; the other 20% served as a test set. The ICU cohort was used as a validation set on the reduced gene sets. Selection was done in a randomly assigned manner. Cross-validation measures were taken by splitting the ER cohort with the function 'KFold' (k = 5) or, where appropriate, cross-validating with a leave-one-out approach due to class imbalances with High severity. Selection and extraction methods were performed to combat overfitting observed when classifying with all DEGs on extreme phenotypes (High + Int. vs. Low) due to a class imbalance with High. We used a recursive feature elimination with cross-validation (RFECV) based on the random forest algorithm from the sklearn library and sequential feature selection (SFS) in the forward and backward direction from the 'mxlextend' (version 0.23.0). [45] However, random noise made the most optimal gene set challenging to determine. In the end, a feature importance based on coefficients with the logistic regression and model-independent metric mutual information were employed by adding the next most important gene per fold based on cross-validated results. Feature extraction via principal component analysis (PCA) was also undertaken on the entire DEG set and reduced sets. Lastly, the final model, logistic regression with L2 regularization, was optimized and validated on the ICU cohort.

All the steps mentioned above are summarized in a flowchart (Figure 1).

Flowchart

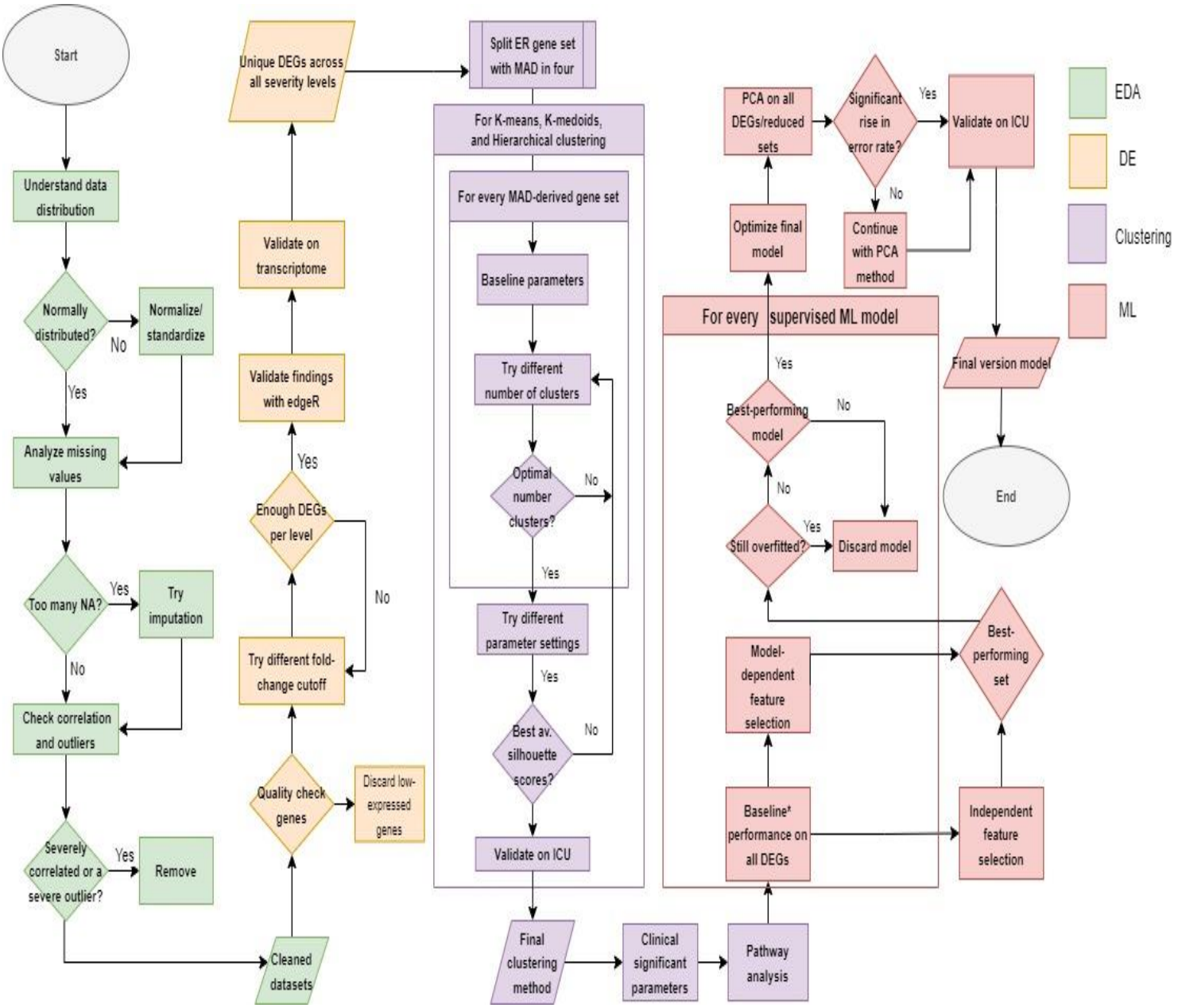


Figure 1—Flowchart regarding the flow of the project. The exploratory data analysis section is highlighted in green. Choices regarding this section are about how to deal with missing values, corrections, and outliers. Thereafter, in yellow, DE analysis mainly regarded choices on the fold-change cutoff value for finding sufficient DEGs per severity level and validation of findings on the transcriptome. We tested different unsupervised clustering (in purple) methods based on multiple gene sets based on MAD (see Materials and Methods) and choosing the best-performing clustering method. Also, using established clusters to extract clinically significant parameters and discover upregulated and downregulated pathways. In red, machine learning was used to predict severity and endotype groups. We tested on different supervised methods and chose the best-performing model, which did not exhibit any signs of overfitting after feature selection.

Results

Discovery of DEGs and associated pathways

We included all 348 patients (ER and ICU) in a differential expression (DE) analysis to establish sepsis severity and mortality gene signatures. However, mortality did not exhibit sufficient DEGs and correlated with severity; therefore, a separation would not add value, and we preferred to use sepsis severity as our primary focus. We compared all groups with each other and extracted unique genes that had sufficient fold changes (1.2) and had a significant adjusted p-value of 0.05. When using the DESeq2, this amounted to 210 unique DEGs out of 1.488 total mitochondria-related genes. To validate our results, we compared our findings to edgeR (Figure 2a) while using the same parameters where possible. We found that both methods label 165 genes as differentially expressed, while edgeR found two and DESeq2 nonoverlapping 45 genes, respectively. Most DEGs were identified by both methods, and some discrepancies are expected since the methods have different calculation methods.

Figure 2b depicts how unstable groups of sepsis severity based on gene expression of found DEGs cluster together by dimension reduction through PCA. We highlight the DEGs found when comparing High and Low in Figure 2c, yielding the highest amount of DEGs of all comparisons (182), meaning that 86,67% of all unique DEGs were also in this comparison. High vs. Intermediate showed 20 and Intermediate vs. Low 54 (see supplementary S2) DEGs, respectively.

Up- and downregulated pathways were identified by comparing unique DEGs per group using Reactome and MSigDB datasets illustrated in Figures 3a and 3b. Comparing High vs. Low severity groups yielded the most significant altered pathways, especially those related to mitochondrial dysfunction, such as downregulated electron transport and mitochondrial translation pathways, as seen in Figure 3a. However, using MSigDB, as shown in Figure 3b, only a few pathways were significantly altered based on High vs. Low. These pathways are related to oxidative phosphorylation, cell-mediated apoptosis, and the breakdown of reactive oxygen species. In contrast to Figure 3a, the High vs. Intermediate comparison did not result in significantly altered pathways.

Figure 2– Discovery of DEGs related to sepsis severity. a) Venn diagram indicating the DEGs found by DESeq2 and validation method edgeR. Most DEGs found by both methods overlap. b) Clustering based on the 210 DEG set found by DESeq2 on sepsis severity status indicates the groups' unstableness. c) Volcano plot of found DEGs (adj. *p-value* of 0.05; absolute FC of 1.2) in comparing High vs. Low sepsis severity groups, yielding a total of 182 DEGs, which makes up 86,67% of the total set.

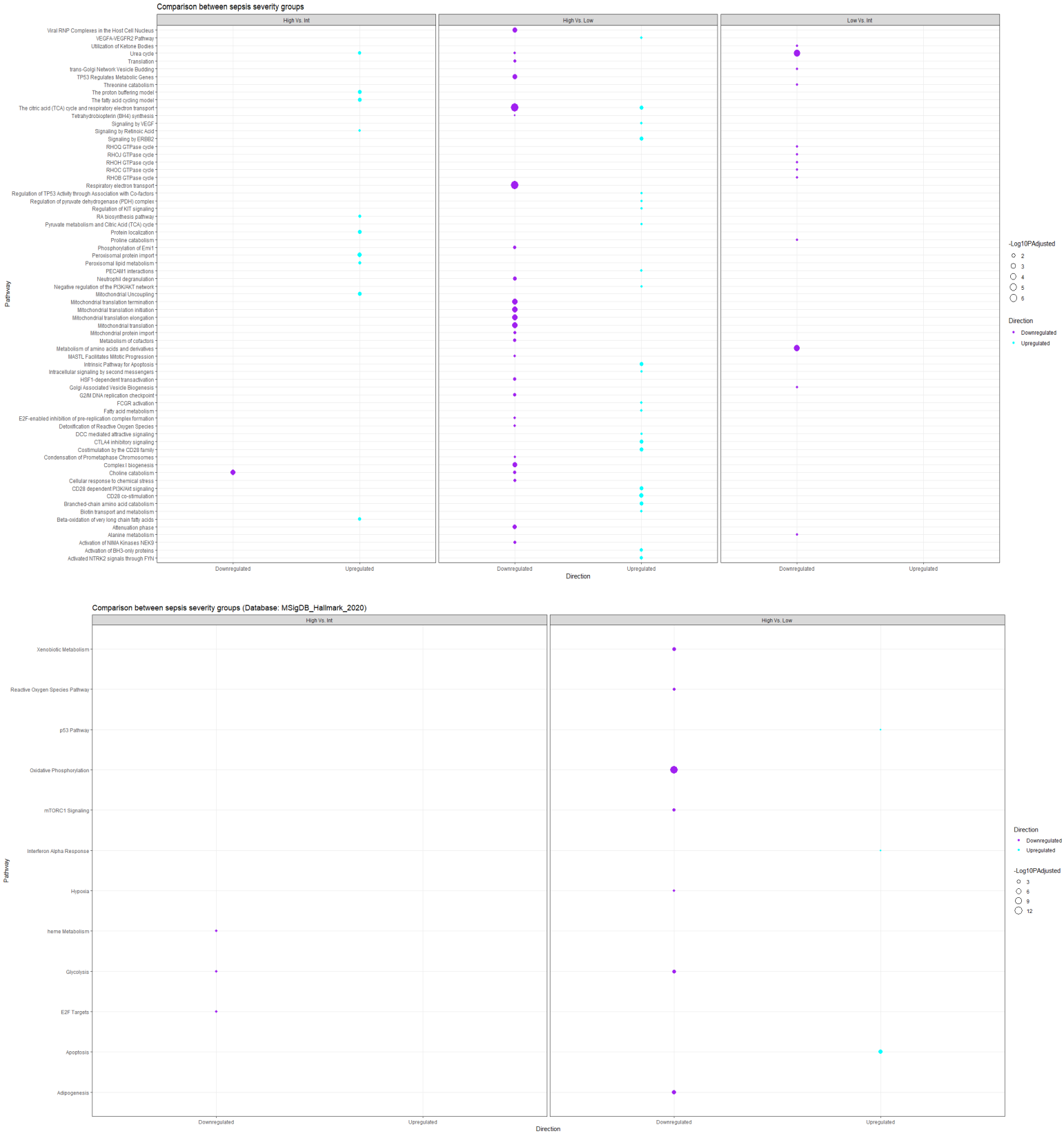


Figure 3– Up- and downregulated pathways based on sepsis severity using all 210 DEGs. a) Reactome database showed many significant up- and downregulated pathways (adjusted *p-value* of 0.05) for every comparison between sepsis severity groups (High, Intermediate, and Low). b) MSigDb's Hallmark 2020 database up- and downregulated pathways. Only sepsis severity groups compared to High had significantly expressed pathways.

Establishing severity-based clusters

The DEGs found in regard to severity groups represent molecular signatures to distinguish between severity levels of sepsis. Leveraging the identification of the DEGs, we proceeded to cluster patients into distinct severity endotypes.

We considered three different clustering methods, namely K-Medoids, K-Means, and hierarchical clustering. After testing on different-sized gene sets based on median absolute deviation (MAD) on the ER cohorts (n = 266), we decided that K-Medoids was the best clustering method out of the three used (Figure 4b, Table 2). K-medoids with the Manhattan distance provided the most stable clusters for the training and test sets regarding DEGs with average silhouette scores of 0,18 and 0,22 (on k = 2), respectively (Figure 4c, d).

After testing on different gene sets based on MAD and three optimization metrics (Table 3), k = 2 proved to be the most optimal number of clusters for all unsupervised clustering methods. In addition, consensus clustering on K-medoids also revealed that 2 clusters were preferred (Figure 4a). When validating on the ICU cohort, we observed comparable results and determined that k = 2 is also the most optimal number of clusters here (See supplementary S3). In addition to clustering primarily on DEGs, we clustered on the entire population for unbiased validation. Clustering on the entire population of mito-genes (1.488) revealed that k = 2 provided the most stable clusters (see supplemental S3).

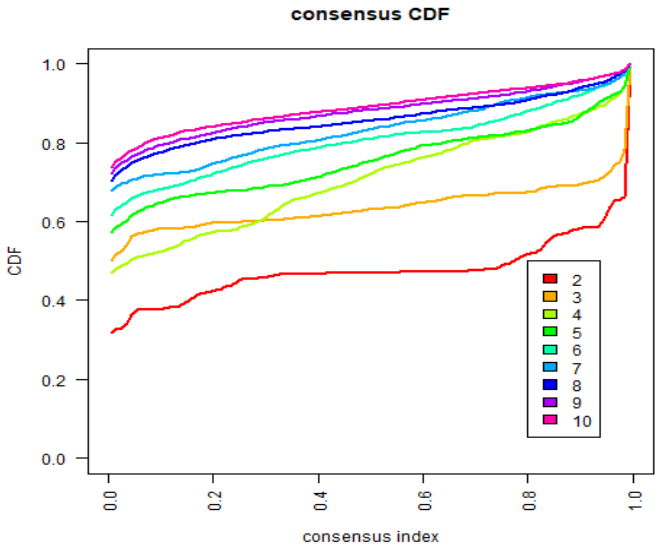
However, K-medoids were not the preferred clustering method for all mito-genes. In contrast, K-Means with Euclidean distance exhibited stable clusters with an average silhouette score of 0,42 out of 1 for k = 2 (see supplemental S3). Nevertheless, when performing on the same parameter settings on DEGs, results were lackluster in comparison with a silhouette score of 0,15 (see supplemental S3). Therefore, K-medoids were still selected as the preferred clustering method.

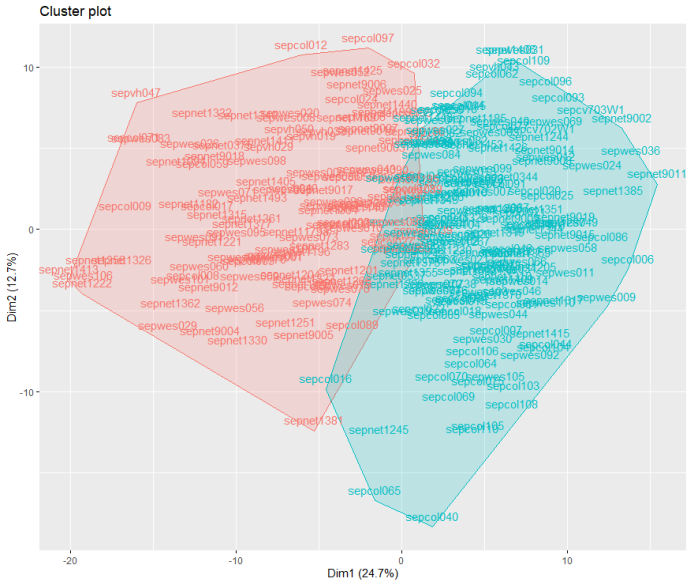
Clustering method	ER	ICU
K-means (Euclidean)	0,16	0,18
Hierarchical (Euclidean + Ward.D2)	0,16	0,14
K-medoids (Euclidean)	0,17	0,19
K-medoids (Manhattan)	0,18	0,22

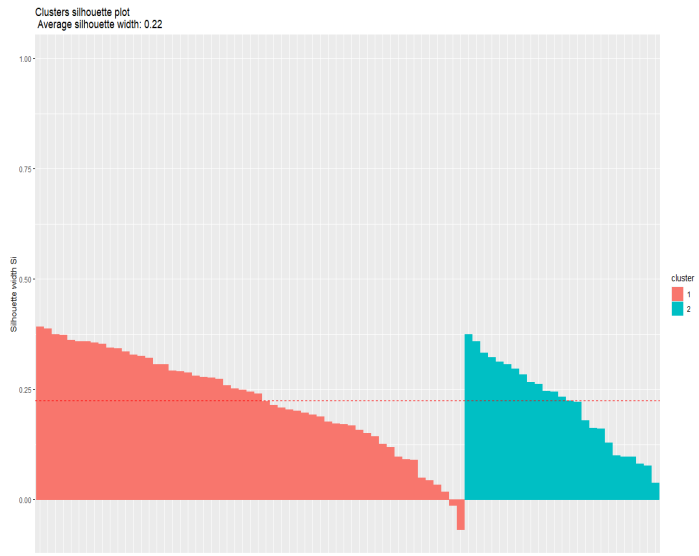
Table 2—All used clustering methods with scoring based on average silhouette scores for k = 2 for ER (training) and ICU (test/validation). PAM had slightly better average silhouette scores for both the training and test/validation and was subsequently chosen as the preferred clustering method.

% of total genes based on MAD	WSS	Silhouette score	Gap score
25%	2-3	2	6
50%	2-3	2	7
75%	2-3	2	7
100%	2-4	2	2

Table 3—Optimal cluster numbers based on the within-cluster sum of squares (WSS), silhouette scores, and gap scores for different ranked-based sets of genes using median absolute deviation (MAD) in batches of 25% for K-medoids. Most indices indicated that 2 clusters for all ER cohorts are the most stable.







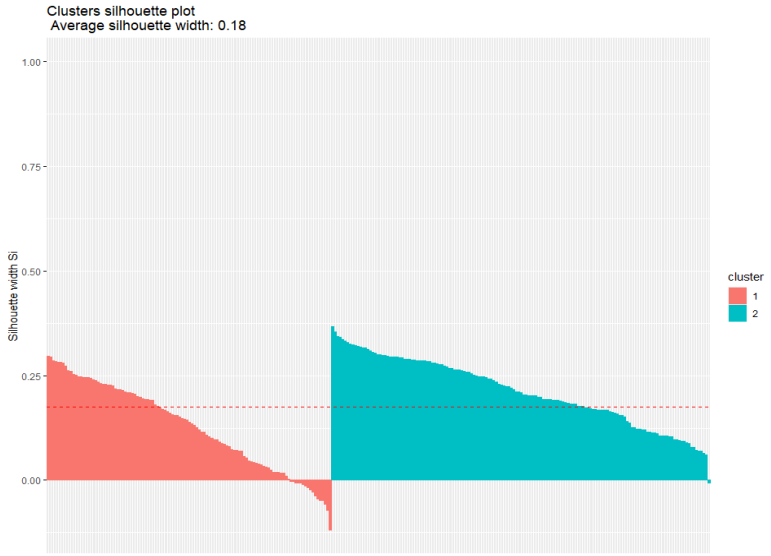


Figure 4–Clusters established based on K-Medoids. a) The Consensus CDF plot indicates two as the most stable number of clusters (red line) for ER cohorts. b) K-medoid clusters with 'ER-severe' (1) in red and 'ER-mild' (2) in blue. c) Cluster stability (K-medoids) based on silhouette scores with 'ICU-severe' (1) in red and 'ICU-mild' (2) in blue. A small piece of cluster one at the edge between both clusters is considered unstable. d) Cluster stability (K-medoids) based on silhouette scores with 'ER-severe' (1) in red and 'ER-mild' (2) in blue. A small piece of cluster one at the edge between both clusters is considered unstable.

Clusters show clinically relevant differences

Establishing two endotypes provided a more meaningful way to categorize sepsis based on mitochondria-related genes. We investigated whether these clusters were distinct from each other in severity by comparing them between each other and healthy controls.

When comparing the established severity-based clusters, we found that 46 out of 267 variables for the entire ER population are considered significant (Table 4). Some significant clinical variables are related to SOFA, namely "First At Ed Sofa" (adj. p-value of 0,000079), "Sepsis Severity" (adj. p-value of 0,0001249), and "Worst within 72h Sofa" (adj. p-value of 0,0004973). In Table 4, we also observed some crucial indicators of mitochondrial dysfunction due to oxidative stress: "First At Ed Lactate" (adj. p-value of 0,000051), "Treatment Oxygen Therapy" (adj. p¬-value of 0.003658), and "Worst Within 72 Lungs Fio2" (adj. p-value of 0.0000015), which indicates an estimation of oxygen content a patient inhales. Evidently, some important molecular metrics that indicate the severity of sepsis were also presented, various metrics regarding immune cell count, whether it is total cell count ("Worst Within Total Cell Count" with adj. p-value of 0,000079) or specific classes of immune cells such as neutrophils ("First at Ed Neutrophil count" with adj. p-value of less than 0,000001).

Significant Variables for Cluster Groups (K-Medoids clustering)

	Cluster Groups		
Parameter	1	2	p.value
Worst Within 72 Heart Rate	100.55 +/- 21.94 (97)	94.54 +/- 18.89 (109)	0.0140381
Age	59.49 +/- 19.02 (115)	53.46 +/- 21.26 (152)	0.0129318
Other Condition Long Term Steroid	100% (115/115)	98.69% (151/153)	0.0118505
At Ed Duration Of Illness Prior To Arrival To Ed	6.17 +/- 18.21 (115)	5.98 +/- 10.11 (150)	0.0108289
Sample Location Colombia	100% (115/115)	98.69% (151/153)	0.0069558
Outcome Icu Admission	100% (115/115)	100% (153/153)	0.0045821
First At Ed Cns Gcs	14.44 +/- 1.78 (108)	14.92 +/- 0.32 (147)	0.0042467
Worst Within 72 Kidney Creatinine	97.58 +/- 81.47 (84)	72.57 +/- 90.82 (65)	0.0034598
Outcome Hospital Admission	100% (115/115)	98.69% (151/153)	0.0032964
First At Ed Coagulation Inr	1.59 +/- 1.52 (37)	1.13 +/- 0.42 (38)	0.0020348
Treatment Fluid Resuscitation	1.01 +/- 0.9 (104)	0.74 +/- 0.6 (143)	0.0014731
Worst Within 72 Body Temperature	37.75 +/- 0.96 (91)	37.12 +/- 3.14 (102)	0.0013458
First At Ed Coagulation Pt	17.9 +/- 18.37 (32)	13.02 +/- 5.66 (35)	0.0012963
First At Ed Eosinophil Count	0.05 +/- 0.08 (94)	0.09 +/- 0.16 (121)	0.0012166
Additonal Test Procalcitonin	11.57 +/- 20.13 (27)	4.07 +/- 14.61 (35)	0.0009666
Treatment Hospital Admission	100% (115/115)	98.69% (151/153)	0.0008462
Outcome Survival Days Organ Failure Free	24.86 +/- 7.93 (115)	27.19 +/- 3.64 (151)	0.0008183
Next 72 Sepsis	100% (115/115)	98.69% (151/153)	0.0008613
Micro Blood Culture Pathogen	98.26% (113/115)	96.08% (147/153)	0.0006294
Worst Within 72 Lactate	2.54 +/- 1.75 (36)	1.33 +/- 1.22 (12)	0.0005735
At Ed Qsofa	1.15 +/- 0.83 (115)	0.81 +/- 0.75 (151)	0.0004973
At Ed Altered Mental State	100% (115/115)	98.69% (151/153)	0.0004495
Sepsis All 72	100% (115/115)	98.69% (151/153)	0.0004573
Treatment Oxygen Therapy	99.13% (114/115)	98.69% (151/153)	0.0003658
Sepsis 3 Excl 1	100% (115/115)	98.69% (151/153)	0.0003586
First At Ed Basophil Count	0.13 +/- 0.93 (94)	0.09 +/- 0.17 (120)	0.0002502
Treatment Icu Admission	100% (115/115)	98.69% (151/153)	0.0001646
Sepsis 3 Bc	98.26% (113/115)	96.08% (147/153)	0.0001530
Sepsis Severity	100% (115/115)	100% (153/153)	0.0001249
First Ed 24 Sepsis	100% (115/115)	98.69% (151/153)	0.0001162
Sepsis 3	100% (115/115)	98.69% (151/153)	0.0001162
Sepsis Org Dys Bc	98.26% (113/115)	96.08% (147/153)	0.0000998
Sepsis 3 Excl 1 2	100% (115/115)	98.69% (151/153)	0.0000775
Treatment Antibiotics Given	100% (115/115)	98.69% (151/153)	0.0000519
First At Ed Liver Billirubin	19.6 +/- 32.7 (99)	13.05 +/- 14.79 (104)	0.0000417
Worst Within 72 Sofa	2.02 +/- 2.83 (115)	0.77 +/- 1.49 (151)	0.0000191
First At Ed Total Cell Count	11.68 +/- 7.45 (115)	8.2 +/- 4.72 (150)	0.0000078
First At Ed Sofa	2.59 +/- 2.17 (115)	1.51 +/- 1.66 (153)	0.0000079
Worst Within 72 Total Cell Count	11.06 +/- 5.92 (104)	7.8 +/- 5.16 (110)	0.0000079
First At Ed Lactate	2.03 +/- 1.47 (84)	1.24 +/- 0.56 (77)	0.0000051
Outcome Hospital Stay Days	9.97 +/- 10.56 (113)	5.65 +/- 6.38 (147)	0.0000041
Worst Within 72 Lungs Fio2	30.9 +/- 16.41 (39)	21.82 +/- 2.67 (51)	0.0000015
Worst Within 72 Neutrophil Count	8.59 +/- 4.75 (64)	5.08 +/- 4.59 (85)	0.0000009
First At Ed Lymphocyte Count	0.9 +/- 0.85 (95)	1.45 +/- 1.57 (126)	0.0000001
First At Ed Neutrophil Count	10.1 +/- 6.98 (98)	5.89 +/- 4.03 (130)	0.0000000
Endotype Name	100% (115/115)	98.69% (151/153)	0.0000000

Table 4–Variables regarding ER patients tested on significance between clusters (*p-value adjusted* 0.05) with Wilcoxon rank-sum test and Chi-Squared test for numeric and categorical, respectively. After that, an overarching post-hoc analysis with Bonferroni was conducted. For both clusters, mean value, standard error, and total available values were included for numeric variables, and the share of available data was compared in percentage.

Cluster ER-mild contained the most severe sepsis cases, as the SOFA scores are higher for qSOFA and the highest measurement of SOFA within the first 72 hours after admission than we observed in cluster two (Figure 5a, d). In addition, biomarkers for mitochondrial dysfunction due to oxidative stress, such as lactate and neutrophil counts, were also higher in cluster ER-severe (Figure 5b, c). There was a difference in mean SOFA scores between both clusters: ER-mild exhibited a mean SOFA score of 0,768 and ER-severe 2,017.

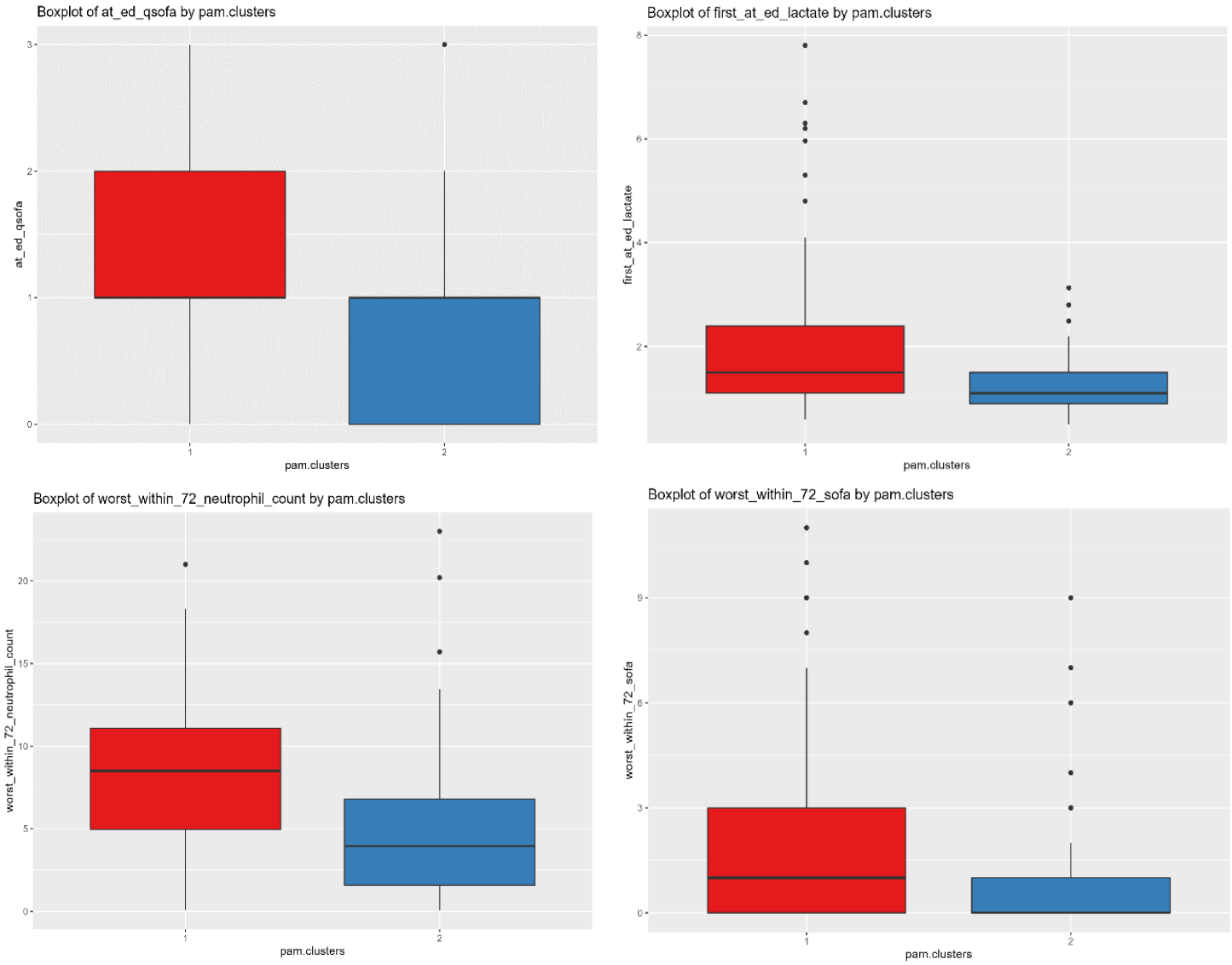
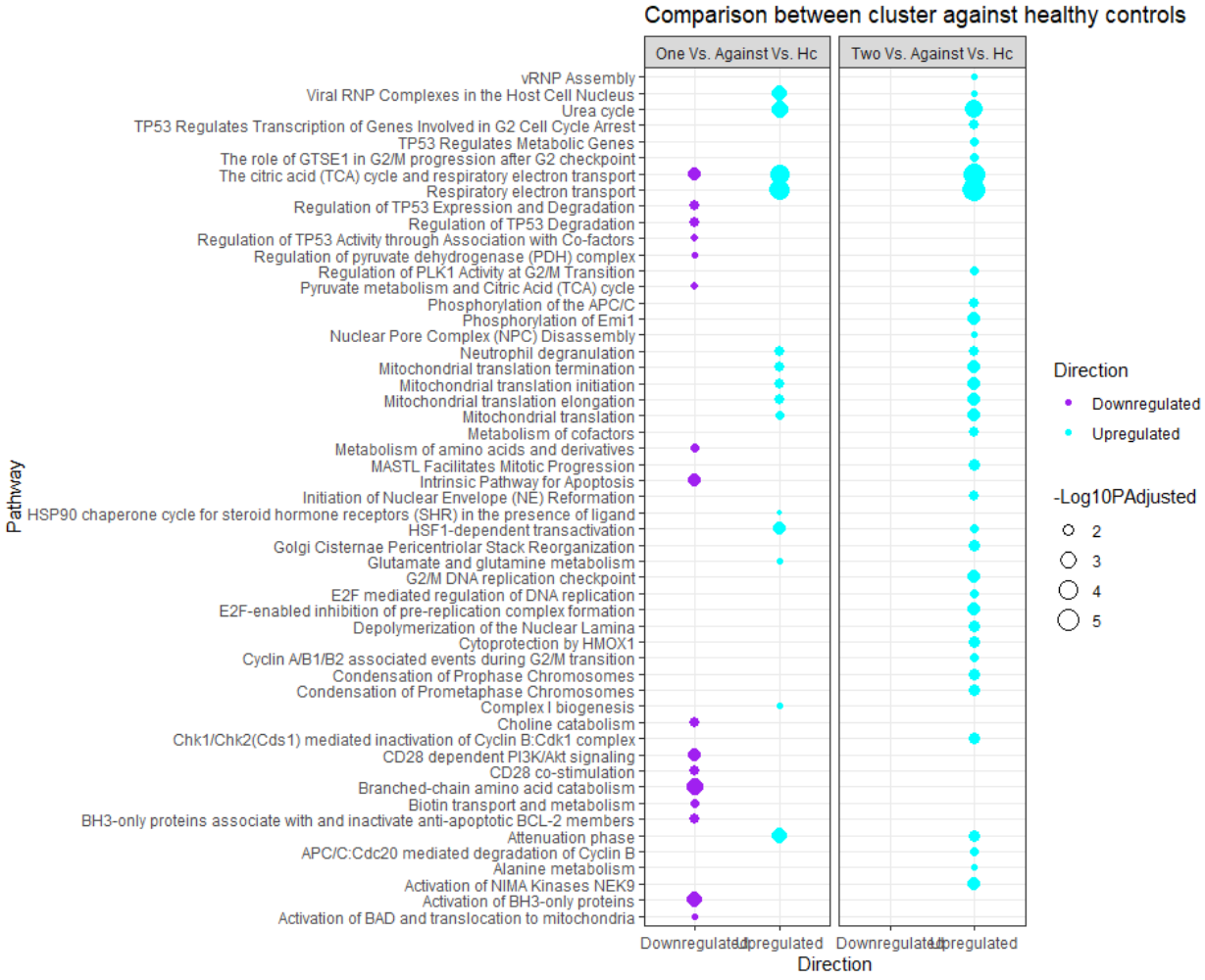


Figure 5—Clinical differences between cluster groups (ER). Cluster 'ER-severe' (1) is depicted in red, and cluster 'ER-mild' (2) in blue. a) Cluster ER-severe (1) showed higher values of qSOFA scores than observed in cluster ER-mild. b) Higher rates of lactate are observed in cluster ER-severe. c) The highest neutrophil count in the first 72 hours after admission was worse in cluster ER-severe. d) Cluster ER-severe exhibited higher SOFA scores than cluster ER-mild.

When comparing up- and downregulated pathways based on log-fold changes between the two established clusters with a healthy control population (n = 44), it is observed that pathways related to mitochondrial and cell electron transportation, mitochondrial translation, elongation, and termination factors, and various immune-related (e.g., neutrophil degranulation and CD28) were upregulated for both cluster ER-severe and ER-mild. Notably, we did not observe any downregulated pathways in the comparison between healthy controls and cluster ER-mild. In contrast, cluster ER-severe exhibited various downregulation, namely in pathways related to apoptosis, regulation of TP53, and respiratory electron transportation. These observations suggested a differential regulatory response in these clusters compared to the healthy control population (Figure 6a). In addition, when comparing gene expression patterns between both clusters, cluster ER-severe contained a higher share of High severity and consequently had elevated SOFA scores than cluster ER-mild (Figure 6b).



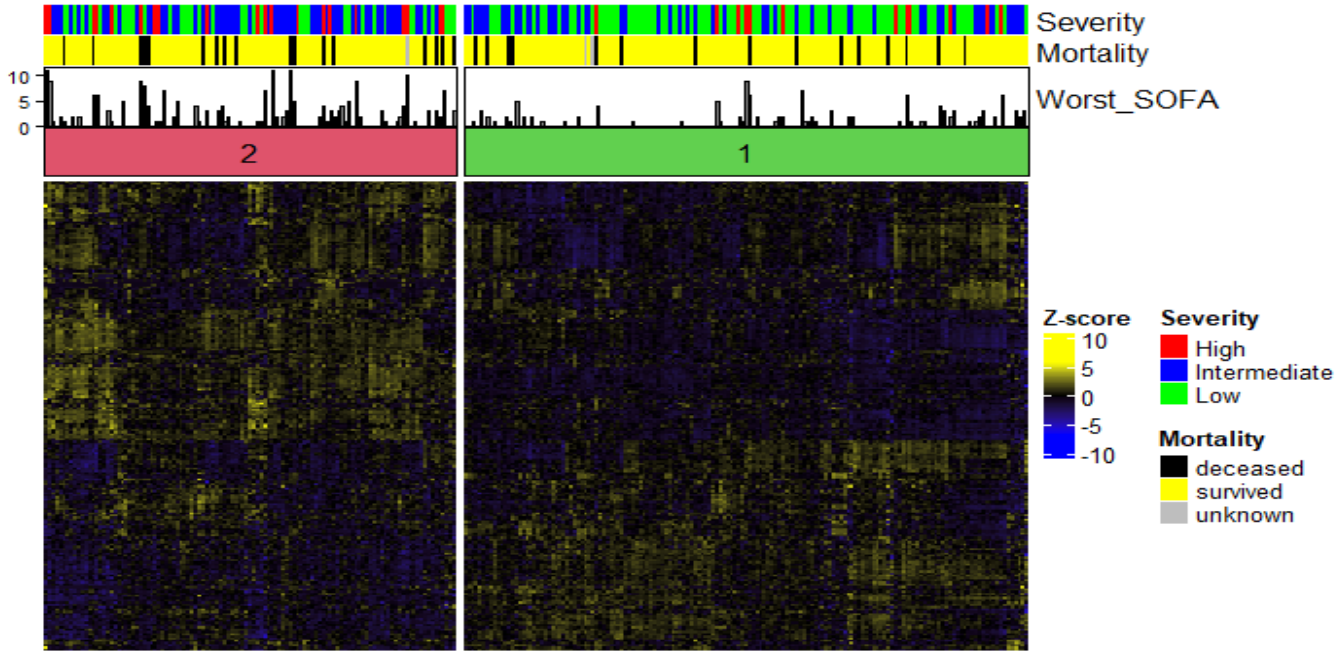


Figure 6—Differences between cluster groups for ER. a) Pathway analysis on cluster groups against a healthy control population of 44 patients. Only significant pathways with a *p*-value less than 0.05 were incorporated. Only the first cluster, in comparison with healthy controls, exhibited downregulated pathways, especially related to apoptosis and TP53. In both clusters, we observed various upregulated pathways, some of which are related to immune cells (neutrophil degranulation), mitochondria-related transcriptional factors, and the respiratory system. b) Heatmap of the two ER-based clusters with gene expression patterns from all 210 DEGs. Cluster ER-severe exhibited higher amounts of High severity (red) than we observed in cluster ER-mild. In addition, the highest measured SOFA scores in the first 72 hours after admission were higher in cluster ER-severe as well.

When validating the results from clustering on the ICU cohort, we observed similar clinically relevant variables were significant to ER (Table 5). For instance, multiple parameters regarding SOFA scores proved to be of significance, i.e., via "ICU Outcome Sofa First," which highlighted the first measurement of SOFA in an ICU setting (adj. p-value of 0.0022) and "ICU Sofa" (adj. p-value of 0.0057) regarding the highest measured SOFA score during admission. We found, as we did for ER cohorts, that neutrophil count and white blood cell measurements ("ICU ANC" and "ICU WBC" with an adj. p-value of >0.0001) are of significance. Additionally, some of the named variables show a clear distinction between the two clusters in Figure 7a-c. Here, the first cluster clearly exhibited higher ANC, WBC, and SOFA measures in contrast to cluster two.

Significant Variables for Cluster Groups (K-Medoids clustering)

	Cluster Groups		
Parameter	1	2	p.value
Rna Quantity	109.96 +/- 117.4 (55)	148.16 +/- 84.24 (25)	0.0049461
First At Ed Sofa	6.53 +/- 4.54 (55)	9.56 +/- 5.33 (25)	0.0057068
Icu Rna Quantity	109.96 +/- 117.4 (55)	148.16 +/- 84.24 (25)	0.0049461
Icu Sofa	6.53 +/- 4.54 (55)	9.56 +/- 5.33 (25)	0.0057068
Icu Outcome Sofa First	6.73 +/- 4.24 (55)	9.96 +/- 4.99 (25)	0.0022916
Icu Apache	20.17 +/- 7.3 (52)	27.21 +/- 10.6 (24)	0.0010500
Icu Outcome Icu Survival	98.18% (54/55)	100% (25/25)	0.0005968
Icu Outcome Icu Mortality	98.18% (54/55)	100% (25/25)	0.0005968
Icu Outcome Survival 28 Day	24.8 +/- 7.47 (54)	18.08 +/- 10.6 (24)	0.0007431
Icu Wbc	9.84 +/- 7.19 (55)	16.12 +/- 6.99 (25)	0.0000088
Icu Anc	7.53 +/- 4.9 (54)	14 +/- 6.07 (25)	0.0000009

Table 5— Variables regarding ICU patients tested on significance between clusters (*p-value* adjusted 0.05) with Wilcoxon rank-sum test and Chi-Squared for numeric and categorical, respectively. After that, an overarching post-hoc analysis with Bonferroni was conducted. For both clusters, mean value, standard error, and total available values are included for numeric variables, and the share of available data is compared to the percentage for categorical data.

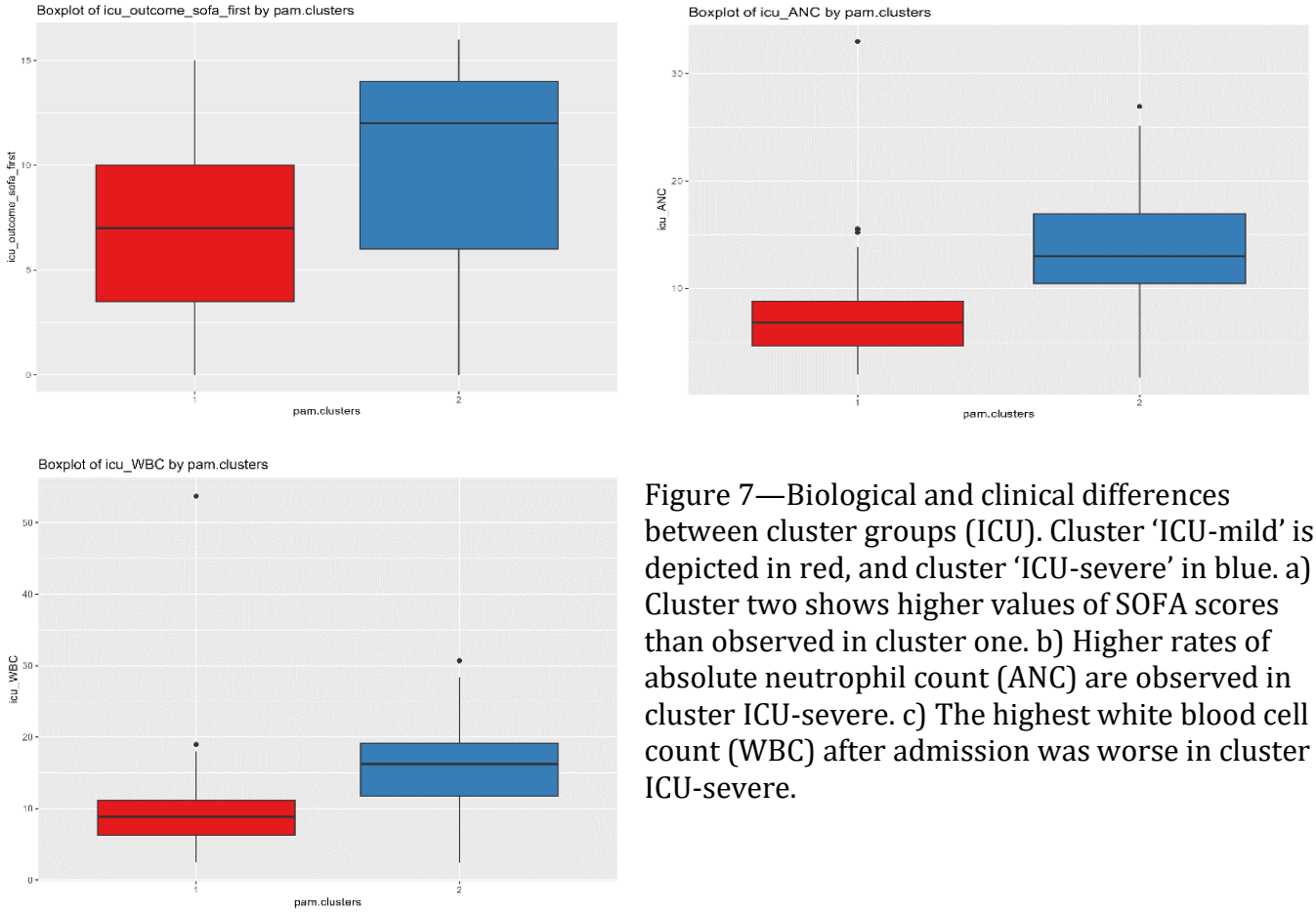
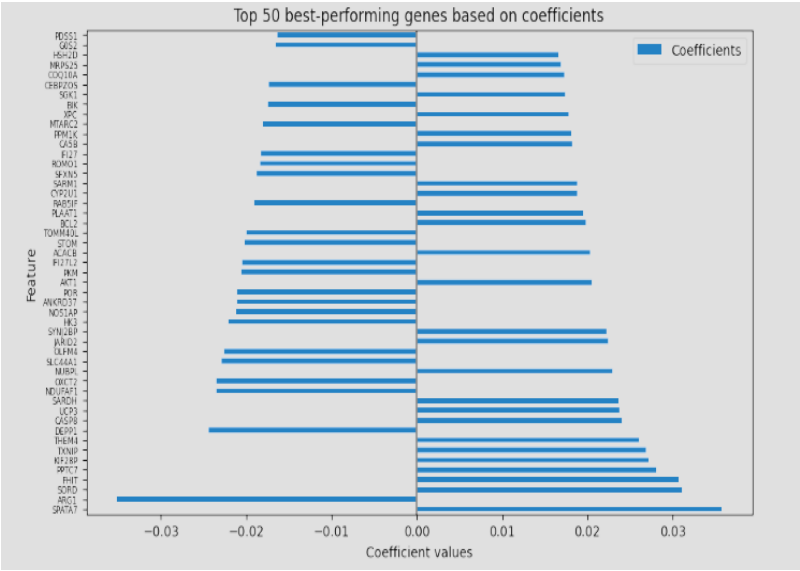


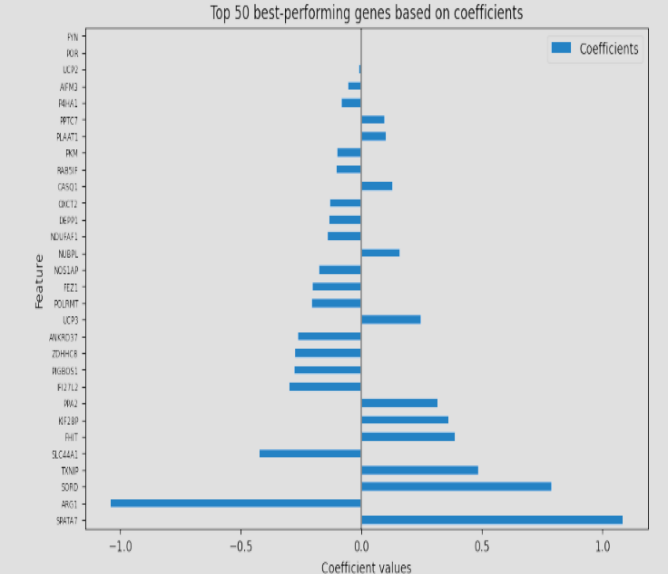
Figure 7—Biological and clinical differences between cluster groups (ICU). Cluster 'ICU-mild' is depicted in red, and cluster 'ICU-severe' in blue. a) Cluster two shows higher values of SOFA scores than observed in cluster one. b) Higher rates of absolute neutrophil count (ANC) are observed in cluster ICU-severe. c) The highest white blood cell count (WBC) after admission was worse in cluster ICU-severe.

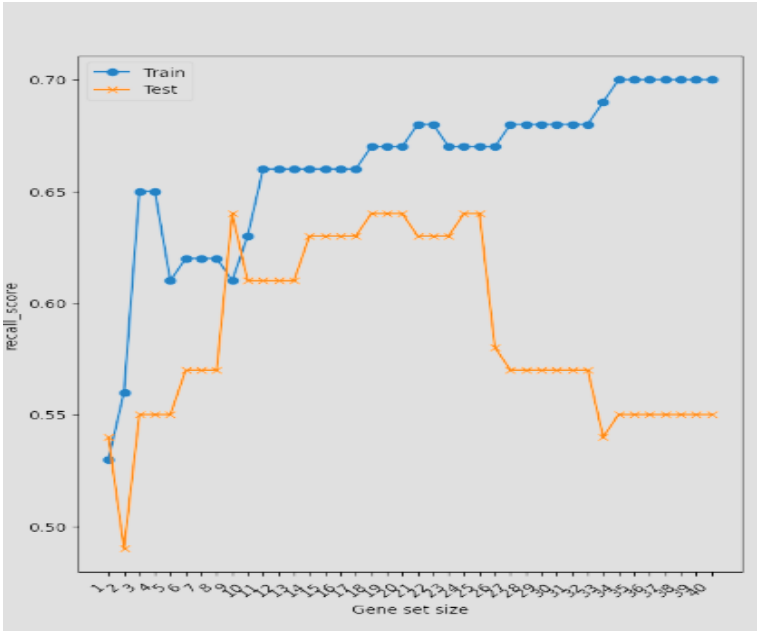
Similar to the observations made in Figure 6, we noticed that a similar pattern existed in the ICU cohort: one cluster exhibited higher SOFA scores than the other cluster (see Appendix A). Namely, the ICU-severe cohort had an average SOFA score of 9,56, and the ICU-mild cohort was 6,53. Additionally, the pathways that were upregulated with the ER clusters were also observed, and in sometimes a larger capacity, in the different ICU-mild and ICU-severe clusters compared to the healthy controls (see Appendix A).

Feature selection

The clusters proved to be distinct from each other in severity and how far mitochondria function had deteriorated. To develop a prognostic tool, we turned to extracting the most important genes that had the best predicting capabilities. Using all 210 DEGs for predicting the cluster number and severity of a patient, we encountered overfitting for all four models (See supplementary S4). Performing various feature selection methods resulted in three gene sets with various sizes. We predicted dependent on the High + Intermediate vs. Low for feature selection. The threshold for mutual information-based feature selection proved somewhat arbitrary, but as seen in Figure 9c, recall scores stayed reasonably consistent between the 9th and 27th highest-ranked genes. In addition, recall scores between the training and test sets were close enough to determine that overfitting was no longer a problem. Therefore, we evaluated this independent feature selection method on all used models for the nine highest-scoring genes. Feature selection based on coefficients from logistic regression, which had acceptable accuracy and recall scores based on RFECV and SFS (see supplementary S4) and is a good candidate for model-based feature selection, were good indicators of gene importance. A logistic regression model using standard parameter settings (L1 or LASSO regularization) found a gene set of four, and an optimized version with L2 or ridge regularization concluded that eleven genes yielded the best recall scores (Figures 9a, b). Gene sets had overlaps between them, especially between four and eleven. However, the nine-gene set only had one overlapping gene with eleven (Figure 9d).







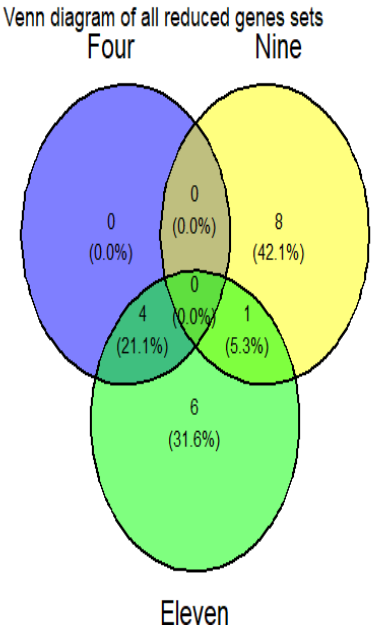


Figure 9—Different methods of feature selection. a) Feature selection through mutual information. The difference between the training (in blue) and test (in orange) scores stays fairly consistent between 0.60 and 0.65 based on recall scores throughout evaluation on a gene set of 9 and 27. Therefore, we selected these nine genes and evaluated all four models for performance. b) Logistic regression using standard parameters (L1 regularization or "LASSO") indicated via coefficients that, when selecting for features on an absolute value of 0.50, four genes were selected. c) Logistic regression using optimized parameters (L1 regularization or "LASSO") indicated via coefficients that, when selecting for features on an absolute value of 0.025, eleven genes were selected. d) Venn diagram with every gene set and how they overlap between the different selection methods.

Training and test of reduced gene sets on ER cohorts

The three reduced gene sets scored fairly consistently for predicting cluster and severity status, as seen in Figure 9a-f. No matter the gene size, AUC scores were continually above 0.85 regarding predicting the cluster number on the ER test set (Figure 10b, d, f). Interestingly, AUC scores for the optimized logistic regression and naïve Bayes were perfect for the nine-gene set (Figure 10f). ROC/AUC scores regarding severity were more varied across the different gene sets. For instance, the eleven-sized gene set had the best scores for logistic regression, with an AUC score of 0.69. The best AUC score belonged to Naïve Bayes, scoring an AUC score of 0.71, also seen in the eleven-sized gene set (Figure 10c). Overall, the nine- and eleven-sized gene sets scored higher than the gene set consisting of four genes.

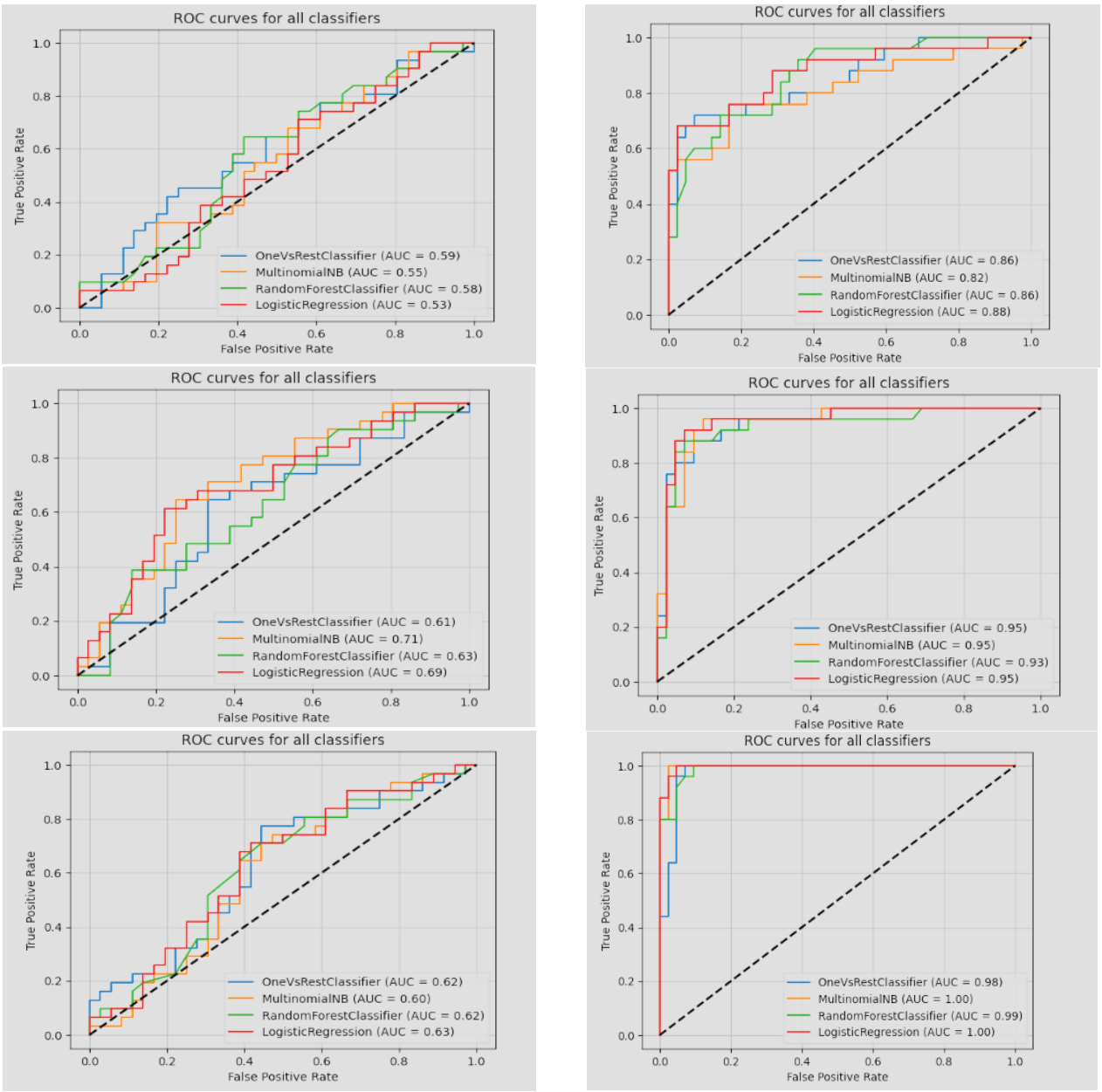


Figure 10—ROC regarding severity and cluster prediction scores of the ER-based test set based on different gene sets. a) ROC of predicting severity based on the four-gene set. b) ROC of predicting cluster based on the four-gene set. c) ROC of predicting severity based on the eleven-gene set. d) ROC of predicting cluster based on the eleven-gene set. e) ROC of predicting severity based on the nine-gene set. f) ROC of predicting cluster based on the nine-gene set.

Validation on ICU cohort

The ICU cohort was used as a validation set of all reduced gene sets and model optimization regarding logistic regression. The results seen in Figure 11a-f highlighted these findings. As we observed for all gene sets, SVM and Random Forest were not able to accurately predict either cluster or severity, collapsing to a mere AUC score of 0.50. However, for logistic regression and Naïve Bayes, AUC and ROC scores were fairly similar to the results seen in the ER test set and sometimes even scored better. For instance, the eleven-sized gene set had, for logistic regression, an AUC score of 0.74, which was also the highest overall score for predicting severity (Figure 11c). In addition, Naïve Bayes' highest score regarding severity was observed in a set consisting of four genes with an AUC of 0.69 (Figure 11a). Both models were perfectly capable of predicting cluster numbers, scoring an AUC of above 0.88 across all gene sets.

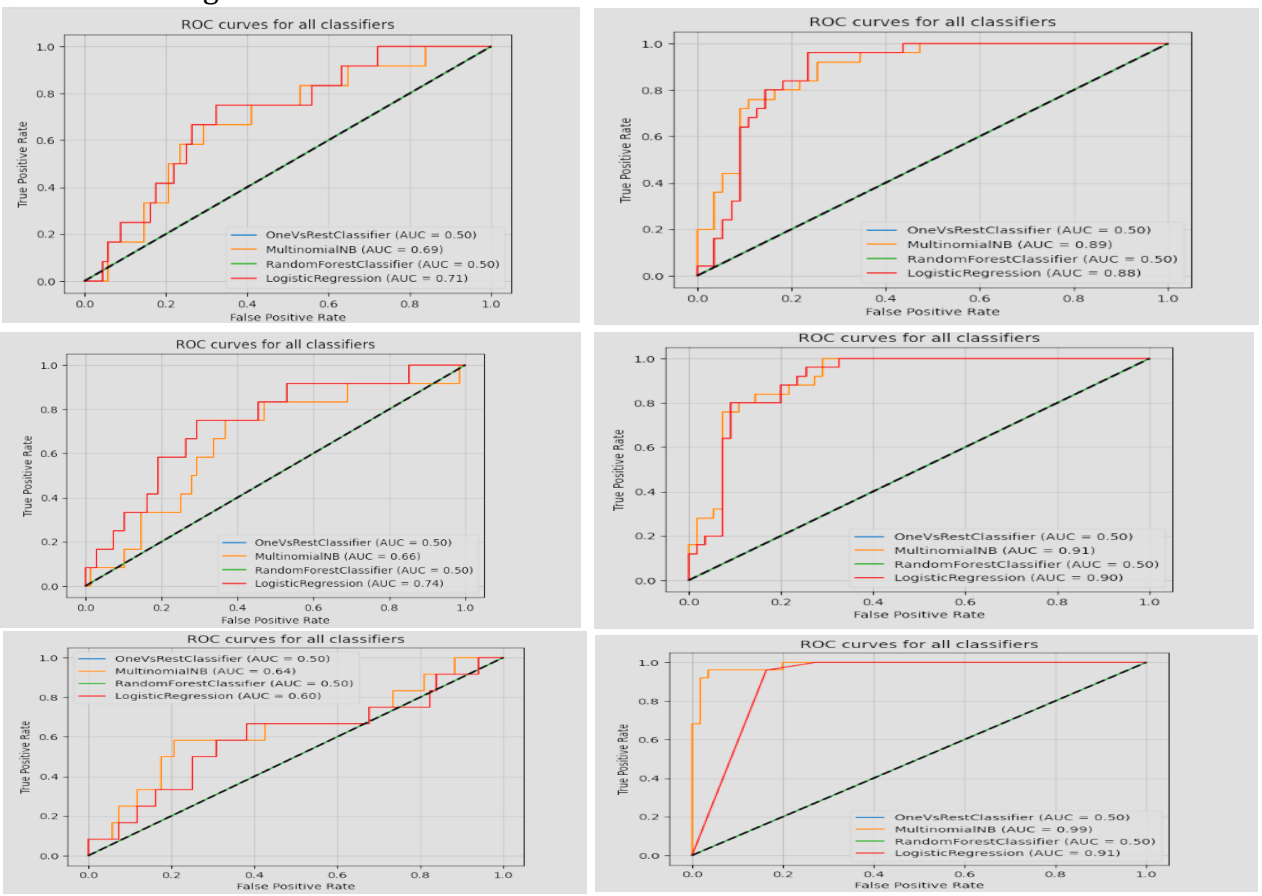


Figure 11—ROC regarding severity and cluster prediction scores of validation set (ICU) based on different gene sets. a) ROC of predicting severity based on the four-gene set. b) ROC of predicting cluster based on the four-gene set. c) ROC of predicting severity based on the eleven gene set. d) ROC of predicting cluster based on the eleven-gene set. e) ROC of predicting severity based on the nine-gene set. f) ROC of predicting cluster based on the nine-gene set.

Final gene set: a zoom-in

After establishing the three reduced gene sets through feature selection and observing prediction capabilities, observing how these genes are expressed by cluster and severity is crucial. Figure 12a-d depicted the gene expressions from the eleven-sized gene set for both ER and ICU cohorts. Here, we observed that some genes are expressed fairly differently for the ER cohorts (Figure 12a, b). For instance, *ARG1* and *TXNIP* are more upregulated for ER-severe in contrast to ER-mild. The opposite can be said for *THEM4* and *FHIT*; these two genes are more downregulated when comparing the two clusters (Figure 12a). Others did not exhibit much change in expression, such as *KIF28P*. Genes that were upregulated also tended to be upregulated in higher severity; ARG1 was upregulated in high severity compared to low severity (Figure 12b). And the reverse was also true. Higher severity onset in the ICU made differences in gene expressions more pronounced. For example, the differences in ARG1 expression between ICU-severe and ICU-mild were higher than in the ER cohort (Figure 12c). This is true for every gene. For instance, we did not observe many differences in expression for PPTC7 for the ER cohort, but in the ICU cohort, the differences were observable. Other gene sets are also included and showed the same observations made for this gene set; whenever a gene was upregulated or downregulated between clusters in the ER, these differences were more pronounced in the ICU cohort. See Appendix B for the four-sized gene set and Appendix C for the nine-sized gene set.

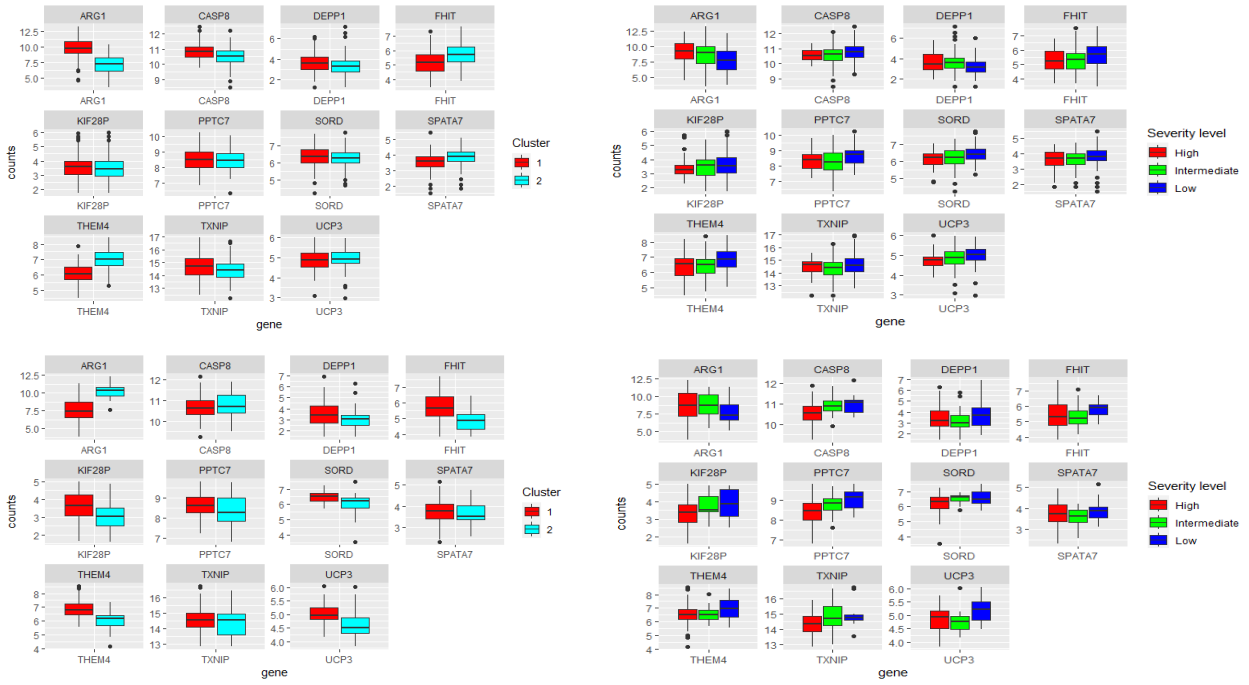


Figure 12—Gene expression patterns of the eleven-sized gene set across ER and ICU cohorts. a) ER cohort gene expression per cluster. b) ER cohort gene expression per severity level. c) ICU cohort gene expression per cluster. d) ICU cohort gene expression per severity level.

Discussion and Conclusion

For this study, we looked at a subset of mitochondrial-related genes to identify potential biomarkers for stratification to build hypotheses for broader studies by focusing on this specific gene subset. The two established clusters show a distinction in gene expression of various biological pathways, especially those related to normal mitochondria function. Some of these pathways were altered due to mitochondrial dysfunction due to oxidative stress, such as apoptosis and respiratory electron transportation, which are commonly seen in septic patients. [13] [46] One of the two clusters (ER-severe and ICU-severe) has higher rates of lactate, higher rates of white bloodcells, higher average SOFA score for both ER and ICU patients, longer admission rates, and a higher overall share of High and Intermediate severity. In addition, the role of neutrophils in sepsis may be a future therapeutic target that significantly differed between both clusters. [47] These observations make them biologically and mechanistically different. Therefore, we termed them severity-based and mitochondria-driven endotypes. Even though we differentiated them between ER and ICU for clarity, both exhibited similar patterns.

Mitochondria are part of highly regulated pathways. Therefore, any changes in gene expression patterns concerning mitochondrial biogenesis due to oxidative stress would be very noticeable. [48] For this reason, we decided to go with the fold-change cutoff value of 1.2; even though there is no well-established fold-change cut-off, 1.5 or 2 is more standard in other similar studies. [49] This way, we identified 210 DEGs, which were used to establish two distinct clusters, providing more genes than the other two mentioned thresholds. However, this way of loose selection criteria was a tradeoff between including DEGs and adding noise within the data and could have affected the stability of clustering. The highest silhouette score observed of 0.18 was with K-medoids using Manhattan distance on the ER set, which is not great but signifies known challenges when clustering based on gene expression data derived from any heterogeneous disease, including sepsis. [50] These challenges include capturing all genetic diversity and the interplay of demographic factors. [50] [81] However, performing on different-sized gene sets via MAD, using various other clustering methods, validating on the ICU cohort, dimension reduction via PCA, and using all mitochondria-related genes did not change these results drastically. Interestingly, most samples were from the Low and Intermediate severity subgroups, indicating that at least severe septic cases were stratified more stable. Additionally, the samples indicated from the Low subclass had relatively higher SOFA scores, almost being identified as Intermediate. From these indications, there is an accurate way of discrimination between both extreme clusters. However, the mentioned shortcomings may limit the scope of septic patients being accurately stratified into one of the established endotypes. For this reason, we cannot foresee an implementation of these endotypes as a prognostic or diagnostic tool at this point in time.

We produced an optimized logistic regression model that can accurately predict the endotype a septic patient belongs to, no matter how large the gene set was and which model we used. We expected this to be the case since the clusters are based on the expression patterns of the DEGs. However, we saw a slight decrease in accuracy when predicting the cluster variable with the reduced gene sets. Feature selection did not yield

any “optimal” subset of mitochondria-related genes that would make a great prediction between the extreme phenotypes (High + Intermediate vs. Low). Having 210 DEGs makes feature selection rather difficult; comparing all gene combinations with different sizes is computably hard to reach, and we observed with, for example, RFECV that performance did not improve much with adding more features at a certain point. Therefore, the main objective was to identify the few genes with the greatest accuracy score, not necessarily the best-performing gene set. For this reason, we performed feature selection on the best-performing model, which resulted in highly specific gene sets, especially for the four- and eleven-sized gene sets. Nevertheless, the one comprising eleven genes out of the three subsets we found had the best sensitivity and specificity outcomes, scoring an AUC of 0,69 on the ER test set and 0,74 on the ICU validation cohort with the optimized logistic regression model. Within clinical studies, AUC scores between 0,70 and 0,80 are considered poor performance and might only be acceptable in a limited number of clinical settings. [51] For the time being, we do not recommend the usage of these gene sets for clinical diagnosis due to the fact that predicting capabilities need to be further improved. For this particular model, we recommend increasing the High-severity population via data augmentation or the addition of independent samples since the class imbalance with Low and Intermediate caused insufficient prediction capabilities.

We did not see any signs of overfitting on the training data on any of the reduced gene sets, which we had observed before feature selection. The difference in performance between the ER and ICU validation cohorts could be attributed to the fact that the ICU had a higher share of more severe septic patients, validating that our model has good generalization, at least in recognizing severe sepsis cases. The overlap in the most critical genes found in the three sets might indicate essential biomarkers and might be worthwhile for future research. Nonetheless, these results show that mitochondria-related genes that are altered in expression due to the onset of sepsis play an essential role in further deterioration and that further research to establish clinical diagnostic tools based on mitochondrial gene expression might be valuable to consider.

Our best-performing gene set consisted of eleven genes, namely SPATA7, ARG1, SORD, FHIT, PPTC7, KIF28P, TXNIP, THEM4, DEPP1, CASP8, and UCP3. Some were present in multiple gene sets, underlining their importance as biomarkers related to mitochondria dysfunction. We briefly discussed the role of each gene concerning mitochondria dysfunction or oxidative stress.

SPATA7 is involved in spermatogenesis and retinal functions, both energy-dependent processes, and is associated with the mitochondrion. [52] However, there is no well-established link between the gene and oxidative stress or sepsis. One study observed, through the effects of bisphenol A (BPA), an increase in oxidative stress levels in the testis and a decrease in SPATA7 gene expression. The authors hypothesize the link between BPA and suppressed expression of SPATA7 but do not connect this to oxidative stress. [53]

ARG1 encodes the protein one of the two arginase isoenzymes and is responsible for the cleavage of L-arginine to form urea and L-ornithine. [54] Arg-1 is upregulated in many immune cells during activation and as does the nitric oxide (NO) production, needs L-arginine. When Arg-1 is upregulated, less L-arginine is available for NO production. [55]

Reduced amounts of NO can impair mitochondria function as its role mainly regulates mitochondria against oxidative stress through its influence on ROS production. [55] [56]

SORD is involved in the polyol pathway and encodes for sorbitol dehydrogenase, which converts sorbitol into fructose. [57] Decreased levels of SORD activity have been linked to the subsequent increased levels of sorbitol to oxidative stress and less available NADPH, disturbing the balance between free NADP+ and NADPH. [58]

PPTC7 showed up in multiple gene sets and is a mitochondrial phosphatase, which prevents ROS from accumulating via regulating antioxidant Q10 through its coenzyme CoQ10. [59] In addition, the gene regulates the process of selective degradation of mitochondria by autophagy, as seen in mice. [60]

KIF28P, also known as KLP6 (Uniprot ID: B7ZC32), regulates the transport of mitochondria through mitochondrial and fusion. [61] It can interact, together with KIFIBα with KIF1-binding protein (KBP) to facilitate the localization of mitochondria in neural cells. [62] Under oxidative stress, kinesins involved in transport, especially those regulated by phosphorylation, may move damaged mitochondria in more significant numbers for reparation or degradation through mitophagy cells in axonal transportation. [62] [63] However, these findings have not been linked to KIF28P/KLP6 specifically but have been put into a broader picture regarding kinesins.

TXNIP plays a role in redox homeostasis and its primary function is to increase the production of ROS, which causes oxidative stress and, ultimately, apoptosis by binding to inflammasome NLRP3. [64] [65] The gene has been a target for personalized medicine in several studies regarding chronic kidney disease, cancer, and dysfunctions affected by the onset of sepsis, underlining its importance in our gene set. [65] [66]

THEM4, also known as CTMP, regulates apoptosis by reducing AKT/PKB signaling, a pathway that is upregulated as a protective response against oxidative stress. [67] THEM4 is upregulated in the event of mitochondrial dysfunction caused by oxidative stress. In addition, THEM4 is a known target for therapeutics in several types of cancers. [68]

DEPP-1, also known as DEPP and C10orf10, contributes to increased ROS levels and regulates autophagy in conjunction with FOXO1 or FOXO3, a mechanism to reduce damage from oxidative stress. [69] [70] Therefore, an increased expression of DEPP was expected.

CASP8 encodes for caspase-8, which plays a role as an initiator in cell apoptosis. It recruits any caspase executioner (either caspase-3, -6, or -7) via cleavage, and after that, the process of cell death starts. [71] Therefore, higher expression patterns of CASP8 are in line with mitochondria dysfunction.

UCP3, a gene encoding a mitochondrial-specific protein, is a selective counter against certain types of excessive ROS production and is expressed higher in the case of oxidative stress. [72] One study showed that overexpression of UCP3 in the presence of 4-HNE leads to a reduced increase in ROS production in older mice. [73] This highlights its importance as a potential biomarker for mitochondria dysfunction.

Our results were similar or better in predicting severity compared to other, more traditional biomarkers. For instance, one exploratory study tried to establish 17 different cytokines as biomarkers. [74] IL-6 and IL-8 had good-scoring AUCs of 0,685 and 0,7000, respectively. Another study looked at procalcitonin, IL-6, and IL-8. These established AUCs of 0,92, 0,75, and 0,71, respectively. [75] However, omics data tends to be better at capturing sepsis heterogeneity and, therefore, has the potential to be a better decision-making tool. [76] In addition, it paves the way for a more personalized approach to medicine.

Our study focused on a subset of genes to discover mitochondria-related biomarkers, and even though we validated the found DEGs with the entire transcriptome, we did not account for cell proportions. This comes with some potential limitations. For instance, overlooking cellular heterogeneity since sepsis can affect cell types differently due to differences in the epigenetics of mitochondria, especially in whole-blood genetics. [77] [7] One such method with promising capabilities is CIBERSORT, a much used deconvolution tool to estimate cell composition based on RNA-Seq data. Its practices have been attributed to the discovery of many new therapeutics. [78] Future research should therefore extend the scope established in this study by incorporating cell proportions and subsequently validate and expand upon our findings. This study laid the groundwork for such an approach while highlighting the importance of mitochondrial dysfunction in discriminating septic patients based on severity.

Only a few studies have been conducted that have tried to establish specific mitochondrial-driven endotypes. However, one study tried identifying mitochondria-related biomarkers and determined, based on consensus clustering, that k = 2 was also the most appropriate number of clusters. [49] However, ER and ICU were not distinguished since it was a case-control study. The authors of the study identified three biomarkers, namely BCKDHB, NDUFB3, and LETMD1, none of which we found regarding the best-performing gene set. The ROC curve for validation had an AUC score of 0,768, whereas we had a score of 0,74 on the ICU validation cohort. Another study tried to establish gene signatures for sepsis for prognostic via single-cell sequencing and found five mitochondrial-related genes that were significantly up-or downregulated. After that, the authors managed to identify two distinct cluster groups, each of which had unique expression levels of immune cells. In no particular order, these five genetic biomarkers were NMR5, MAOA, COX7b, PPM1K, and TTC19. [79] Predicting capabilities vary since the authors constructed an over-time prognosis tool for 7-, 14-, and 28-day periods. Nevertheless, most ROC curves had an AUC score between 0,65 and 0,75. Again, no gene overlaps with our DEG set. The existence of at least two endotypes or subgroups in sepsis based on gene expression patterns of mitochondria is in line with other studies, and prediction capabilities were reasonably comparable. The essential predictors differ in size and have no overlap between the studies. Reasons for these observed differences could be the implementation and goals of the named studies. First, our study focused on establishing mitochondria-related biomarkers based on an already established severity measurement, whereas the others did not. Second, both studies extracted DEGs by comparing healthy controls with septic patients; we compared various levels of sepsis severity, not against healthy samples. Lastly, both studies used stricter thresholds regarding fold changes (FC of | 2 |), whereas we were more interested in

detecting smaller changes in gene expression. This resulted in 210 DEGs to base the cluster on and 3 and 5 for the other two studies, respectively. These differences underline the goals of each study; namely, these authors wanted to establish endotypes and explore biomarkers to distinguish between healthy and septic patients. Our goals were more aligned to explore biomarkers to stratify patients into different severity-based endotypes.

Sepsis is a multifaceted condition, with mitochondrial dysfunction only representing one aspect of this complex clinical syndrome. Other heterogeneous characteristics of sepsis come from immune responses, different levels of inflammation, divergent speeds of deterioration, and organ dysfunction. They should be considered for an accurate clinical implementation, which our model lacks. [50] Therefore, focusing solely on RNA-Seq data, or even on gene expression in general, in a clinical setting would not sufficiently address all challenges of the heterogeneous nature of sepsis. [81] [50] Employing multimodal methods could enhance the assessment of septic patients by integrating clinical data relatable to mitochondria dysfunction, such as lactate levels, other types of omics data, elevated levels of ROS, cytochrome c, and cardiolipin levels. [80] Other studies followed a more immune-related gene approach and found, in some cases, better prediction scores than solely focusing on mitochondria-related biomarkers. The study conducted by Baquir and coresearchers is an ideal example; based on a 40-gene set, the prediction capabilities are somewhat better than our model. Predicting sepsis severity signatures achieved an AUC score of 0,75 compared to 0,74 on validation cohorts, respectfully. [9] Therefore, another multimodal prediction model, in addition to being trained on mitochondria-related genes and clinical attributes, the inclusion of immune-related attributes such as neutrophil count and various omics biomarkers, could be the next step in refining this model as a tool for prognostic purposes.

Conclusion and future work

The goal of this project was to establish severity-driven endotypes based on the gene expressions of mitochondria-related genes. We can hereby conclude that this goal has been achieved. We produced one optimized logistic regression model with ridge regularization that can predict the extreme phenotypes of SOFA-based sepsis severity with acceptable scores and accurately the established mitochondria-driven endotype cluster based on a novel eleven-sized gene set. However, because of the ongoing challenges of sepsis heterogeneity, the lack of excellent-scoring AUCs in our results, and the challenges of prognostic capabilities regarding mitochondria-related genes, assisting in clinical decision-making is not feasible and requires more research.

Future work can adhere to establishing how mitochondria-related genes, especially in the context of the reduced gene sets, behave in the entire cell population. This could be undertaken as a validation study or building upon the work done in this study. Also, model-building efforts could include other clinical mitochondria-related parameters to improve prediction regarding mitochondria dysfunction as a proxy for determining sepsis severity. In addition, including other nonmitochondrial-related biomarkers, especially in conjunction with the eleven-sized gene set, could address the heterogeneity of sepsis further, improve the sensitivity and specificity of early sepsis recognition, and eventually pave the way for the development of more personalized medicine.

References

[1] M. Singer *et al.*, "The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)," *JAMA*, vol. 315, no. 8, pp. 801–810, Feb. 2016, doi: 10.1001/jama.2016.0287.

[2] R. Luhr, Y. Cao, B. Söderquist, and S. Cajander, "Trends in sepsis mortality over time in randomised sepsis trials: A systematic literature review and meta-analysis of mortality in the control arm, 2002–2016," *Critical Care*, vol. 23, no. 1, p. 241, 2019, doi: 10.1186/s13054-019-2528-0.

[3] E. K. Stevenson, A. R. Rubenstein, G. T. Radin, R. S. Wiener, and A. J. Walkey, "Two decades of mortality trends among patients with severe sepsis: A comparative meta-analysis*," *Critical Care Medicine*, vol. 42, no. 3, pp. 625–631, 2014, doi: 10.1097/CCM.0000000000000026.

[4] K. E. Rudd *et al.*, "Global, regional, and national sepsis incidence and mortality, 1990–2017: Analysis for the global burden of disease study," *The Lancet*, vol. 395, no. 10219, pp. 200–211, Jan. 2020, doi: 10.1016/s0140-6736(19)32989-7.

[5] J. Hajj, N. Blaine, J. Salavaci, and D. Jacoby, "The "centrality of sepsis": A review on incidence, mortality, and cost of care," *Healthcare*, vol. 6, no. 3, p. 90, Jul. 2018, doi: https://doi.org/10.3390/healthcare6030090.

[6] C. J. Paoli, "Epidemiology and costs of sepsis in the united states-an analysis based on timing of diagnosis and severity level," vol. 46, no. 12, pp. 1889–1897, Jan. 2018, doi: 10.1097/CCM.0000000000003342.

[7] A. Leligdowicz and M. A. Matthay, "Heterogeneity in sepsis: New biological evidence with clinical applications," *Critical Care*, vol. 23, no. 1, Mar. 2019, doi: 10.1186/s13054-019-2372-2.[

8] T. E. Sweeney *et al.*, "Unsupervised analysis of transcriptomics in bacterial sepsis across multiple datasets reveals three robust clusters," *Critical Care Medicine*, vol. 46, no. 6, pp. 915–925, Jun. 2018, doi: 10.1097/ccm.0000000000003084.

[9] A. Baghela *et al.*, "Predicting sepsis severity at first clinical presentation: The role of endotypes and mechanistic signatures," *eBioMedicine*, vol. 75, p. 103776, Jan. 2022, doi: 10.1016/j.ebiom.2021.103776.

[10] S. Lambden, P. F. Laterre, M. M. Levy, and B. Francois, "The SOFA score—development, utility and challenges of accurate assessment in clinical trials," *Critical Care*, vol. 23, no. 1, Nov. 2019, doi: 10.1186/s13054-019-2663-7.

[11] S. Maitra, A. Som, and S. Bhattacharjee, "Accuracy of quick sequential organ failure assessment (qSOFA) score and systemic inflammatory response syndrome (SIRS) criteria for predicting mortality in hospitalized patients with suspected infection: A meta-analysis of observational studies," *Clinical Microbiology and Infection*, vol. 24, no. 11, pp. 1123–1129, Nov. 2018, doi: 10.1016/j.cmi.2018.03.032.

[12] M. P. Murphy, "How mitochondria produce reactive oxygen species," *Biochemical Journal*, vol. 417, no. 1, pp. 1–13, Dec. 2008, doi: 10.1042/bj20081386.

[13] M. Ott, V. Gogvadze, S. Orrenius, and B. Zhivotovsky, "Mitochondria, oxidative stress and cell death," *Apoptosis*, vol. 12, no. 5, pp. 913–922, Feb. 2007, doi: 10.1007/s10495-007-0756-2.

[14] K. Nakahira, S. Hisata, and A. M. K. Choi, "The roles of mitochondrial damage-associated molecular patterns in diseases," *Antioxidants & Redox Signaling*, vol. 23, no. 17, pp. 1329–1350, Dec. 2015, doi: 10.1089/ars.2015.6407.

[15] M. Singer, "The role of mitochondrial dysfunction in sepsis-induced multi-organ failure," *Virulence*, vol. 5, no. 1, pp. 66–72, Nov. 2013, doi: 10.4161/viru.26907.

[16] B. S. Star, E. C. van der Slikke, A. van Buiten, R. H. Henning, and H. R. Bouma, "The novel compound SUL-138 counteracts endothelial cell and kidney dysfunction in sepsis by preserving mitochondrial function," *International Journal of Molecular Sciences*, vol. 24, no. 7, p. 6330, Mar. 2023, doi: 10.3390/ijms24076330.

[17] S. Andrews, "FastQC: A quality control tool for high throughput sequence data."

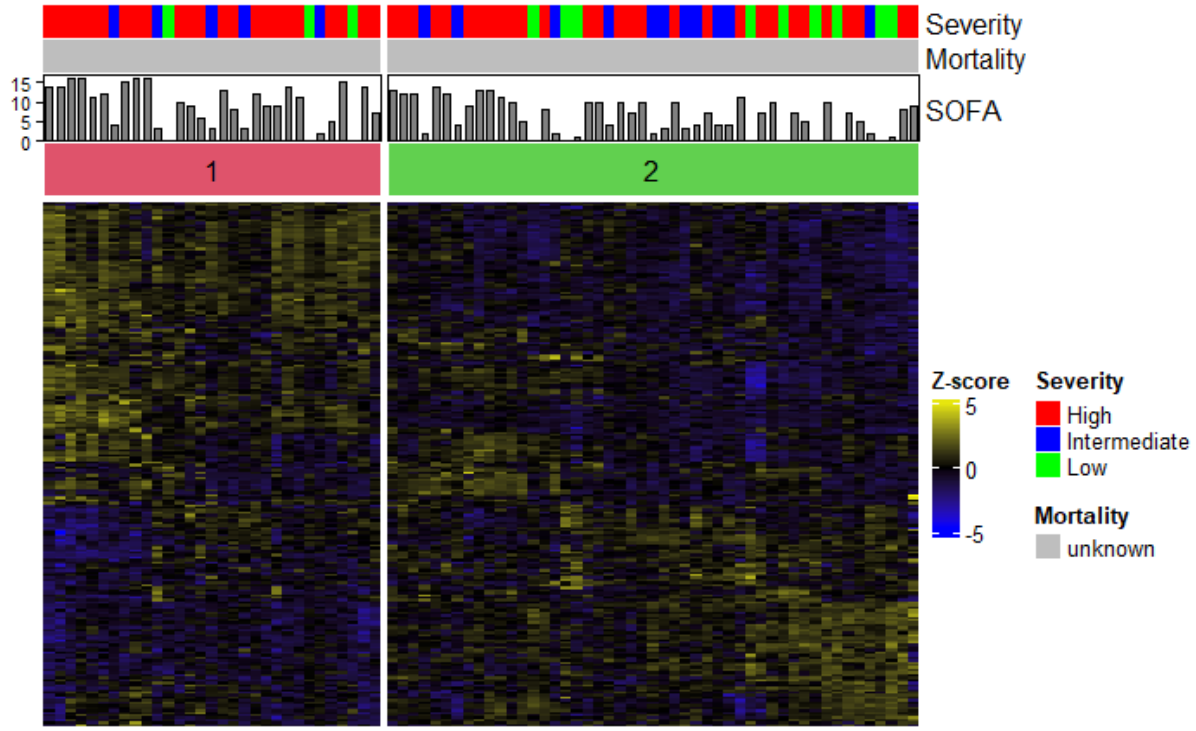
http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, 2010.
[18] P. Ewels, M. Magnusson, S. Lundin, and M. Käller, "MultiQC: Summarize analysis results for multiple tools and samples in a single report," *Bioinformatics*, vol. 32, no. 19, pp. 3047–3048, Jun. 2016, doi: 10.1093/bioinformatics/btw354.
[19] A. Dobin and T. R. Gingeras, "Mapping RNA-seq reads with STAR," *Current Protocols in Bioinformatics*, vol. 51, no. 1, Sep. 2015, doi: 10.1002/0471250953.bi1114s51.
[20] S. Anders, P. T. Pyl, and W. Huber, "HTSeq—a python framework to work with high-throughput sequencing data," *Bioinformatics*, vol. 31, no. 2, pp. 166–169, Sep. 2014, doi: 10.1093/bioinformatics/btu638.
[21] R Core Team, *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2023. Available: https://www.R-project.org
[22] H. Wickham, R. François, L. Henry, K. Müller, and D. Vaughan, *Dplyr: A grammar of data manipulation*. 2023. Available: https://dplyr.tidyverse.org
[23] H. Wickham *et al.*, "Welcome to the tidyverse," *Journal of Open Source Software*, vol. 4, no. 43, p. 1686, Nov. 2019, doi: 10.21105/joss.01686.
[24] M. Kuhn, "Building predictive models inRUsing thecaretPackage," *Journal of Statistical Software*, vol. 28, no. 5, 2008, doi: 10.18637/jss.v028.i05.
[25] H. Wickham, *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York, 2016. Available: https://ggplot2.tidyverse.org
[26] D. Scheper, "Unraveling the role of mitochondria dysfunction in early sepsis." https://www.bitbucket.org, 2023.
[27] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2," *Genome Biology*, vol. 15, no. 12, Dec. 2014, doi: 10.1186/s13059-014-0550-8.
[28] Y. Chen, A. T. L. Lun, and G. K. Smyth, "From reads to genes to pathways: Differential expression analysis of RNA-seq experiments using rsubread and the edgeR quasi-likelihood pipeline," *F1000Research*, vol. 5, p. 1438, Aug. 2016, doi: 10.12688/f1000research.8987.2.
[29] G. Yu and Q.-Y. He, "ReactomePA: An r/bioconductor package for reactome pathway analysis and visualization," *Molecular BioSystems*, vol. 12, no. 2, pp. 477–479, 2016, doi: 10.1039/c5mb00663e.
[30] M. V. Kuleshov *et al.*, "Enrichr: A comprehensive gene set enrichment analysis web server 2016 update," *Nucleic Acids Research*, vol. 44, no. W1, pp. W90–W97, May 2016, doi: 10.1093/nar/gkw377.
[31] Jeffrey T. Leek <Jtleek@Gmail.Com>, W. Evan Johnson<Wej@Bu.Edu>, Hilary S. Parker <Hiparker@Jhsph.Edu>, Elana J.Fertig <Ejfertig@Jhmi.Edu>, Andrew E. Jaffe <Ajaffe@Jhsph.Edu>, John D. Storey <Jstorey@Princeton.Edu>, Yuqing Zhang<Zhangyuqing.Pkusms@Gmail.Com>, Leonardo Collado Torres<Lcollado@Jhu.Edu>, "Sva." Bioconductor, 2017. doi: 10.18129/B9.BIOC.SVA.
[32] T. Galili, "Dendextend: An r package for visualizing, adjusting and comparing trees of hierarchical clustering," *Bioinformatics*, vol. 31, no. 22, pp. 3718–3720, Jul. 2015, doi: 10.1093/bioinformatics/btv428.
[33] D. M. Wilkerson and N. D. Hayes, "ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking," *Bioinformatics*, vol. 26, no. 12, pp. 1572–1573, 2010.
[34] A. Kassambara and F. Mundt, *Factoextra: Extract and visualize the results of multivariate data analyses*. 2020. Available: https://CRAN.R-project.org/package=factoextra
[35] M. Charrad, N. Ghazzali, V. Boiteau, and A. Niknafs, "NbClust: An r package for determining the relevant number of clusters in a data set," *Journal of Statistical Software*, vol. 61, no. 6, pp. 1–36, 2014, Available: https://www.jstatsoft.org/v61/i06/
[36] Z. Gu, R. Eils, and M. Schlesner, "Complex heatmaps reveal patterns and correlations in multidimensional genomic data," *Bioinformatics*, vol. 32, no. 18, pp. 2847–2849, May 2016, doi:

10.1093/bioinformatics/btw313.
[37] Python Software Foundation, "Python (version 3.11.12)." https://www.python.org/, 2023.
[38] T. Kluyver *et al.*, "Jupyter notebooks – a publishing format for reproducible computational workflows." IOS Press, pp. 87–90, 2016.
[39] W. McKinney *et al.*, "Data structures for statistical computing in python," in *Proceedings of the 9th python in science conference*, Austin, TX, 2010, pp. 51–56.
[40] C. R. Harris *et al.*, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020, doi: 10.1038/s41586-020-2649-2.
[41] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.
[42] M. Waskom *et al.*, "Mwaskom/seaborn: v0.8.1 (september 2017)." Zenodo, Sep. 2017. doi: 10.5281/zenodo.883859.
[43] J. D. Hunter, "Matplotlib: A 2D graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007, doi: 10.1109/mcse.2007.55.
[44] Reiichiro Nakano, "Reiinakano/scikit-plot: v0.2.1." Zenodo, 2017. doi: 10.5281/ZENODO.293191.
[45] S. Raschka, "MLxtend: Providing machine learning and data science utilities and extensions to python's scientific computing stack," *Journal of Open Source Software*, vol. 3, no. 24, p. 638, Apr. 2018, doi: 10.21105/joss.00638.
[46] H. Cui, Y. Kong, and H. Zhang, "Oxidative stress, mitochondrial dysfunction, and aging," *Journal of Signal Transduction*, vol. 2012, pp. 1–13, Oct. 2012, doi: 10.1155/2012/646354.
[47] X. Shen, K. Cao, Y. Zhao, and J. Du, "Targeting neutrophils in sepsis: From mechanism to translation," *Frontiers in Pharmacology*, vol. 12, Apr. 2021, doi: 10.3389/fphar.2021.644270.
[48] C. Ploumi, I. Daskalaki, and N. Tavernarakis, "Mitochondrial biogenesis and clearance: A balancing act," *The FEBS Journal*, vol. 284, no. 2, pp. 183–195, Aug. 2016, doi: 10.1111/febs.13820.
[49] Q. Shu, H. She, X. Chen, L. Zhong, J. Zhu, and L. Fang, "Identification and experimental validation of mitochondria-related genes biomarkers associated with immune infiltration for sepsis," *Frontiers in Immunology*, vol. 14, May 2023, doi: 10.3389/fimmu.2023.1184126.
[50] A. A. Woodward, R. J. Urbanowicz, A. C. Naj, and J. H. Moore, "Genetic heterogeneity: Challenges, impacts, and methods through an associative lens," *Genetic Epidemiology*, vol. 46, no. 8, pp. 555–571, Aug. 2022, doi: 10.1002/gepi.22497.
[51] C. E. Metz, "Basic principles of ROC analysis," *Seminars in Nuclear Medicine*, vol. 8, no. 4, pp. 283–298, Oct. 1978, doi: 10.1016/s0001-2998(78)80014-2.
[52] GeneCards, "SPATA7 gene - spermatogenesis associated 7." https://www.genecards.org/cgi-bin/carddisp.pl?gene=SPATA7, Accessed in 2024.
[53] A. H. Kamel, M. A. Foaud, and H. M. Moussa, "The adverse effects of bisphenol a on male albino rats," *The Journal of Basic and Applied Zoology*, vol. 79, no. 1, Jan. 2018, doi: 10.1186/s41936-018-0015-9.
[54] M. Munder, "Arginase: An emerging key player in the mammalian immune system," *British Journal of Pharmacology*, vol. 158, no. 3, pp. 638–651, Oct. 2009, doi: 10.1111/j.1476-5381.2009.00291.x.
[55] C. J. Darcy *et al.*, "Increased plasma arginase activity in human sepsis: Association with increased circulating neutrophils," *Clinical Chemistry and Laboratory Medicine*, vol. 52, no. 4, Jan. 2014, doi: 10.1515/cclm-2013-0698.
[56] K. Wijnands, T. Castermans, M. Hommen, D. Meesters, and M. Poeze, "Arginine and citrulline and the immune response in sepsis," *Nutrients*, vol. 7, no. 3, pp. 1426–1463, Feb. 2015, doi: 10.3390/nu7031426.
[57] A. Y. W. LEE and S. S. M. CHUNG, "Contributions of polyol pathway to oxidative stress in diabetic cataract," *The FASEB Journal*, vol. 13, no. 1, pp. 23–30, Jan. 1999, doi: 10.1096/fasebj.13.1.23.

[58] P. A. Barnett, R. G. González, L. T. Chylack, and H.-M. Cheng, "The effect of oxidation on sorbitol pathway kinetics," *Diabetes*, vol. 35, no. 4, pp. 426–432, Apr. 1986, doi: 10.2337/diab.35.4.426.
[59] I. González-Mariscal *et al.*, "The mitochondrial phosphatase PPTC7 orchestrates mitochondrial metabolism regulating coenzyme Q10 biosynthesis," *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, vol. 1859, no. 11, pp. 1235–1248, Nov. 2018, doi: 10.1016/j.bbabio.2018.09.369.
[60] N. M. Niemi *et al.*, "PPTC7 maintains mitochondrial protein content by suppressing receptor-mediated mitophagy," *Nature Communications*, vol. 14, no. 1, Oct. 2023, doi: 10.1038/s41467-023-42069-w.
[61] K. Tanaka, Y. Sugiura, R. Ichishita, K. Mihara, and T. Oka, "KLP6: A newly identified kinesin that regulates the morphology and transport of mitochondria in neuronal cells," *Journal of Cell Science*, vol. 124, no. 14, pp. 2457–2465, Jul. 2011, doi: 10.1242/jcs.086470.
[62] N. Hirokawa, Y. Noda, Y. Tanaka, and S. Niwa, "Kinesin superfamily motor proteins and intracellular transport," *Nature Reviews Molecular Cell Biology*, vol. 10, no. 10, pp. 682–696, Oct. 2009, doi: 10.1038/nrm2774.
[63] C. Fang, D. Bourdette, and G. Banker, "Oxidative stress inhibits axonal transport: Implications for neurodegenerative diseases," *Molecular Neurodegeneration*, vol. 7, no. 1, Jun. 2012, doi: 10.1186/1750-1326-7-29.
[64] R. Zhou, A. Tardivel, B. Thorens, I. Choi, and J. Tschopp, "Thioredoxin-interacting protein links oxidative stress to inflammasome activation," *Nature Immunology*, vol. 11, no. 2, pp. 136–140, Dec. 2009, doi: 10.1038/ni.1831.
[65] C. Yang, W. Xia, X. Liu, J. Lin, and A. Wu, "Role of TXNIP/NLRP3 in sepsis-induced myocardial dysfunction," *International Journal of Molecular Medicine*, Jun. 2019, doi: 10.3892/ijmm.2019.4232.
[66] J. Y. Lim, "Thioredoxin and thioredoxin-interacting protein as prognostic markers for gastric cancer recurrence," *World Journal of Gastroenterology*, vol. 18, no. 39, p. 5581, 2012, doi: 10.3748/wjg.v18.i39.5581.
[67] A. Parcellier *et al.*, "Carboxy-terminal modulator protein (CTMP) is a mitochondrial protein that sensitizes cells to apoptosis," *Cellular Signalling*, vol. 21, no. 4, pp. 639–650, Apr. 2009, doi: 10.1016/j.cellsig.2009.01.016.
[68] M. M. Hill and B. A. Hemmings, "Inhibition of protein kinase b/akt," *Pharmacology & Therapeutics*, vol. 93, no. 2–3, pp. 243–251, Feb. 2002, doi: 10.1016/s0163-7258(02)00193-6.
[69] S. Salcher, M. Hermann, U. Kiechl-Kohlendorfer, M. J. Ausserlechner, and P. Obexer, "C10ORF10/DEPP-mediated ROS accumulation is a critical modulator of FOXO3-induced autophagy," *Molecular Cancer*, vol. 16, no. 1, May 2017, doi: 10.1186/s12943-017-0661-4.
[70] M. W. Stepp, R. J. Folz, J. Yu, and I. N. Zelko, "The c10orf10 gene product is a new link between oxidative stress and autophagy," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1843, no. 6, pp. 1076–1088, Jun. 2014, doi: 10.1016/j.bbamcr.2014.02.003.
[71] D. R. McIlwain, T. Berger, and T. W. Mak, "Caspase functions in cell death and disease," *Cold Spring Harbor Perspectives in Biology*, vol. 7, no. 4, p. a026716, Apr. 2015, doi: 10.1101/cshperspect.a026716.
[72] E. Barreiro *et al.*, "UCP3 overexpression neutralizes oxidative stress rather than nitrosative stress in mouse myotubes," *FEBS Letters*, vol. 583, no. 2, pp. 350–356, Dec. 2008, doi: 10.1016/j.febslet.2008.12.023.
[73] M. Nabben *et al.*, "The effect of UCP3 overexpression on mitochondrial ROS production in skeletal muscle of young versus aged mice," *FEBS Letters*, vol. 582, no. 30, pp. 4147–4152, Nov. 2008, doi: 10.1016/j.febslet.2008.11.016.
[74] F. A. Bozza *et al.*, "Cytokine profiles as markers of disease severity in sepsis: A multiplex analysis," *Critical Care*, vol. 11, no. 2, p. R49, 2007, doi: 10.1186/cc5783.
[75] S. HARBARTH *et al.*, "Diagnostic value of procalcitonin, interleukin-6, and interleukin-8 in

critically ill patients admitted with suspected sepsis," *American Journal of Respiratory and Critical Care Medicine*, vol. 164, no. 3, pp. 396–402, Aug. 2001, doi: 10.1164/ajrccm.164.3.2009052.
[76] T. Itenov, D. Murray, and J. Jensen, "Sepsis: Personalized medicine utilizing 'omic' technologies—a paradigm shift?" *Healthcare*, vol. 6, no. 3, p. 111, Sep. 2018, doi: 10.3390/healthcare6030111.
[77] A. E. Jaffe and R. A. Irizarry, "Accounting for cellular heterogeneity is critical in epigenome-wide association studies," *Genome Biology*, vol. 15, no. 2, p. R31, 2014, doi: 10.1186/gb-2014-15-2-r31.
[78] A. M. Newman *et al.*, "Robust enumeration of cell subsets from tissue expression profiles," *Nature Methods*, vol. 12, no. 5, pp. 453–457, Mar. 2015, doi: 10.1038/nmeth.3337.
[79] Q. Shu *et al.*, "Construction and validation of a mitochondria-associated genes prognostic signature and immune microenvironment characteristic of sepsis," *International Immunopharmacology*, vol. 126, p. 111275, Jan. 2024, doi: 10.1016/j.intimp.2023.111275.
[80] J. N. Acosta, G. J. Falcone, P. Rajpurkar, and E. J. Topol, "Multimodal biomedical AI," *Nature Medicine*, vol. 28, no. 9, pp. 1773–1784, Sep. 2022, doi: 10.1038/s41591-022-01981-2.
[81] Tsimberidou, A-M., "Initiative for Molecular Profiling and Advanced Cancer Therapy and challenges in implementation of precision medicine," *Elsevier BV*, vol. 41, p. 176-181, May 2017, doi: http://dx.doi.org/10.1016/j.currproblcancer.2017.02.002

Appendix A: Differences in ICU cohort



[Healthy control ICU figure is missing]

Figure 13— Differences between cluster groups for ICU. a) Pathway analysis on cluster groups against a healthy control population of 44 patients. Only significant pathways with a *p*-value less than 0.05 were incorporated. Only the first cluster, in comparison with healthy controls, exhibited downregulated pathways, especially related to apoptosis and TP53. In both clusters, we observed various upregulated pathways, some of which are related to immune cells (neutrophil degranulation), mitochondria-related transcriptional factors, and the respiratory system. b) Heatmap of the two ICY-based clusters with gene expression patterns from all 210 DEGs. Cluster ICU-severe exhibited higher amounts of High severity (red) than we observed in cluster ICU-mild. In addition, the highest measured SOFA scores in the first 72 hours after admission were higher in cluster ICU-severe as well.

Appendix B: Differences in gene expression (four-sized gene set)

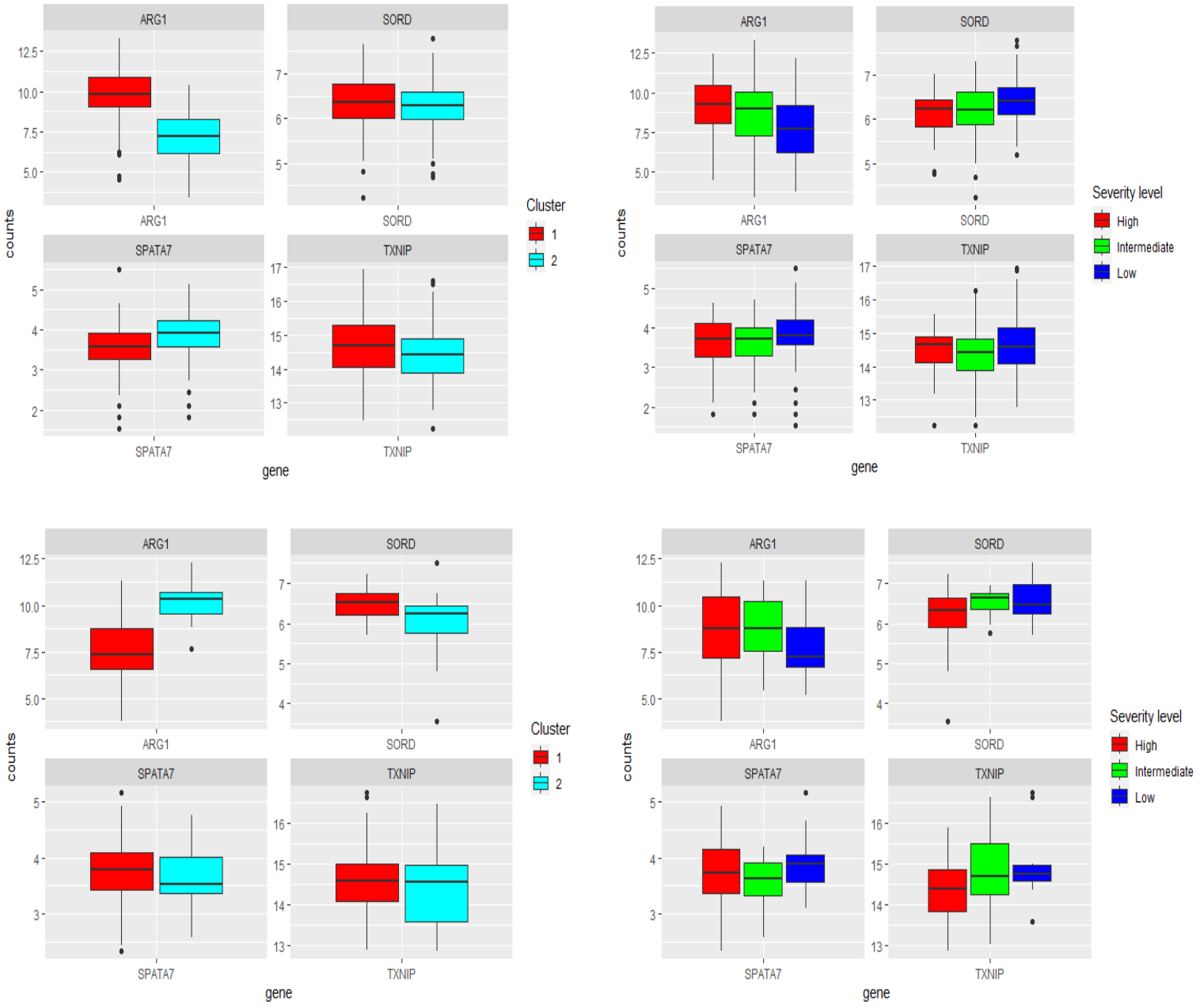


Figure 14—Gene expression patterns of the four-sized gene set across ER and ICU cohorts. a) ER cohort gene expression per cluster. b) ER cohort gene expression per severity level. c) ICU cohort gene expression per cluster. d) ICU cohort gene expression per severity level.

Appendix C: Differences in gene expression (nine-sized gene set)

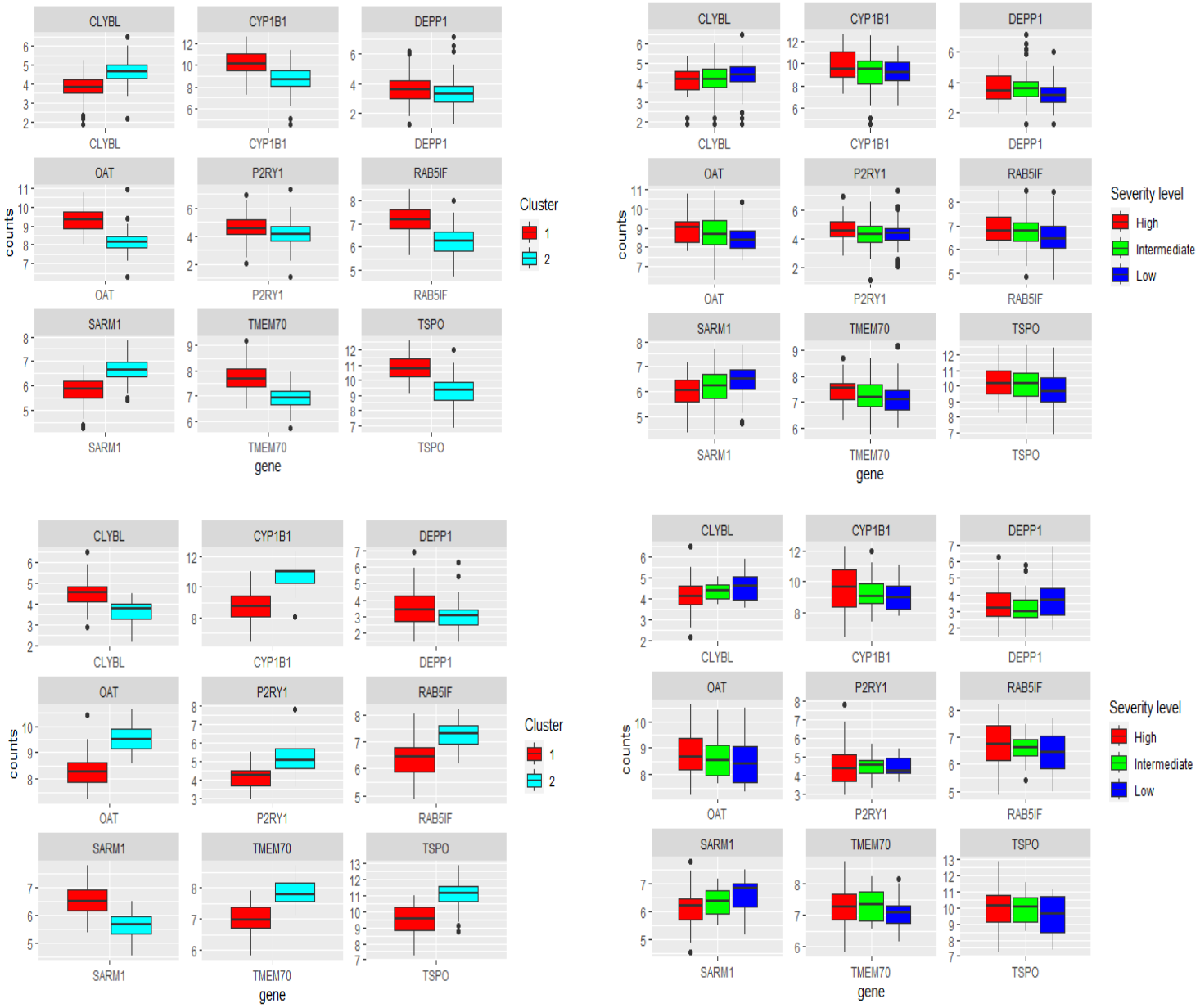


Figure 15—Gene expression patterns of the nine-sized gene set across ER and ICU cohorts. a) ER cohort gene expression per cluster. b) ER cohort gene expression per severity level. c) ICU cohort gene expression per cluster. d) ICU cohort gene expression per severity level.

Appendix D: Dutch translation of abstract

Sepsis is een grote bijdrager aan sterfte in ziekenhuizen. Dit komt voornamelijk doordat sepsis zich op vele manieren kan voordoen. Mitochondriën die beschadigd zijn geraakt door hevige immuunreacties, kunnen sepsis verergeren. Het analyzeren van genen gerelateerd aan mitochondriën kan een vroegtijdige indicatie geven over de gesteldheid van een patiënt en helpen met het opstellen van een behandelplan.

De genprofielen van 348 patiënten met sepsis—afkomstig van vier spoedafdelingen en één intensivecareafdeling–zijn gebruikt en vergeleken met 44 gezonde individuën. Klinisch relevante genprofielen zijn totstandgekomen door het toepassen van verschillende vormen van kunstmatige intelligentie.

Genen die anders tot expressie komen (*DEGs*), zijn geïdentificeerd doormiddel van het vergelijken van groepen patiënten met varierende ernst van sepsis. Deze ernst wordt bepaald door een zogeheten SOFA-score. Er zijn twee unieke groepen totstandgekomen door het gebruik van de *DEGs* op basis van patiënten afkomstig uit de vier spoedafdelingen. Deze groepen worden ook wel endotypen genoemd. De twee endotypen zijn vergeleken met de groep gezonde individuën en gevalideerd op de patiënten van de intensivecare. Doormiddel van selectie op relevante genen, zijn er drie verschillende sets van genen ontdekt die op een accurate wijze de endotypegroep en ernst van sepsis kunnen voorspellen. Met name een set van elf genen had een goed accuratie bij het voorspellen van de endotypegroep (85%) en de ernst van sepsis (76%).

De genprofielen en endotypen die zijn ontdekt, geven de essentiele rol van mitochondriën in het tijdig herkennen van sepsis aan. Toekomstig onderzoek moet gericht zijn op het ontwikkelen van een multimodaal prognostische tool voor een verbetering in stratificatie van patiënten.