Feature selection for the accurate prediction of septic and cardiogenic shock ICU mortality in the acute phase

Alexander Aushev[1◑], Vicent Ribas Ripoll[2◑], Alfredo Vellido[1◑], Federico Aletti[3] Bernardo Bollen Pinto[4,5], Karim Bendjelid[4,5], Antoine Herpain[6], Emiel Hendrik Post[6], Eduardo Romay Medina [7], Ricard Ferrer [8,9], Giuseppe Baselli [10]

1 Intelligent Data Engineering and Artificial Intelligence (IDEAI) Research Center, Computer Science, Universitat Politècnica de Catalunya, Barcelona, Spain.
2 Eurecat, Centre Tecnològic de Catalunya, eHealth, Data Analytics in Omics, Barcelona, Spain.
3 Department of Bioengineering, University of California San Diego, La Jolla, CA, USA
4 Hemodynamic Research Group, Université de Genève, Geneva, Switzerland.
5 Anesthesiology Department, Hôpitaux Universitaires de Genève, Geneva, Switzerland.
6 Department of Intensive Care, Hôpital Erasme - Université Libre de Bruxelles, Brussels, Belgium.
7 Hospital Universitari Mutua de Terrassa, Terrassa, Spain.
8 Critical Care Department, Vall d'Hebron University Hospital, Barcelona, Spain.
9 Shock, Organ Dysfunction and Resuscitation (SODIR) Research Group, Vall d' Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain.
10 Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano - Piazza Leonardo da Vinci, Milan, Italy.

◑These authors contributed equally to this work.
* vicent.ribas@eurecat.org

Abstract

Circulatory shock is a life-threatening disease that accounts for around one-third of all admissions to intensive care units (ICU). It requires immediate treatment, which is why the development of tools for planning therapeutic interventions is required to deal with shock in the critical care environment. In this study, the ShockOmics European project original database is used to extract attributes capable of predicting mortality due to shock at the ICU. Missing data imputation techniques and machine learning models were used, followed by feature selection from different data subsets. Selected features were later used to build Bayesian Networks, revealing causal relationships between features and ICU outcome. The main result is a subset of predictive features that includes well-known indicators such as SOFA and APACHE II, but also less commonly considered ones related to cardiovascular function such as Left Atrial Dilation, Inferior Vena Cava Distensibility Index, or shock treatment (Norepinephrine). Importantly, certain selected features are shown to be most predictive at certain time-steps. This means that, as shock progresses, different attributes could be prioritized. Clinical traits obtained at 24h. from ICU admission are shown to accurately predict cardiogenic and septic shock mortality, suggesting that relevant life-saving decisions could be made shortly after ICU admission.

Introduction

Shock (or circulatory shock) is a life-threatening medical condition that requires immediate treatment. It is prevalent at the Intensive Care Unit (ICU) and a major concern in Critical Care in general. This condition occurs when the organs and tissues of the body do not receive enough blood and, as a result, cells see their oxygen and nutrients supply restricted so that organs become damaged. Low blood pressure, fast pulse rate, rapid breathing and fever are amongst the most common symptoms [28]. The outcome of an individual case depends on the stage of shock, the underlying condition and the general medical state of the patient [28].

Four types of shock are commonly defined: hypovolaemic shock (e.g. hemorrhagic shock), cardiogenic shock, distributive shock (e.g. septic shock) and obstructive shock. The mortality rate of the condition remains very high and depends on its type. It is quantified at 30% for septic shock (according to the new definition of such condition [29]) and between 35.9% and 64.7% for in-hospital mortality in patients with cardiogenic shock (depending on the stage, [30]). Usually, shock is treated in the ICU, where it is carefully monitored. As a result, large amounts of data are produced for each patient.

Machine Learning (ML) techniques can help find a list of attributes to predict the outcome of shock patients from clinical data (i.e. data routinely monitored in the ICU). However, there is a large amount of attributes that characterize each patient, ranging from base pathology information, to lab procedures or monitoring data and treatment, to name a few. As a result, data are likely to be very high dimensional and, therefore, feature selection (FS) methods are often required to reduce data dimensionality and the associated variability. The combination of statistics and ML may provide an appropriate framework to retrieve new knowledge from data, as well as causal relationships between different features from high-dimensional datasets [31–33].

This study analyzes a database gathered within the *ShockOmics* research project. It includes two types of shock: septic and cardiogenic. Septic shock is the most common, and it is caused by infection. The second most frequent form of shock is the cardiogenic shock and it is a result of a heart failure. The multiscale nature of the *ShockOmics* database allows to search for attributes that are related to mortality induced by these two types of shock.

When it comes to the detection of circulatory shock, physicians and therapists largely depend on a combination of clinical, hemodynamic and biochemical signs. For a treatment to be chosen, prompt identification of shock is necessary. Appropriate treatment is based on a good understanding of the physiological mechanisms behind the condition. The management of shock is challenging and patient survival is highly dependent on the timely administration of the appropriate treatment. Normally, it requires the administration of vasoactive drugs and fluid resuscitation. This treatment is usually given to counteract the other conditions that go along with shock (e.g. hypotension/hemodynamic instability, inflammation and multiple organ failure (MOF)). The fact that they have similar symptoms makes addressing the causes of shock very difficult. As a result, decision making at the onset of the condition is not a trivial task and, therefore, it is important to design fast, reliable and interpretable tools to plan therapeutic interventions to prevent mortality, or any irreversible consequences caused by shock.

The study of the pathophysiology of shock and its management has been an active area of research in ML applications for Critical Care. For example, fuzzy decision support systems (DDS) for the management of post-surgical cardiac intensive care patients have been described in [4], while the issue of rule generation for septic shock patients has been addressed in [5] and [6], the former together with an artificial neural network (ANN). The expert system described in [8] is more closely related to the general management of sepsis and it was designed for the detection of pathogens and the

prescription of antibiotics during the first hours of evolution. Ross and co-workers [9] later derived a system of ordinary differential equations together with an ANN model of inflammation and septic shock. Other studies have deployed ANNs for the study of sepsis. They include [7], who presented a clinical study examining SIRS and Multiple organ dysfunction syndrome (MODS) in the ICU after cardiac and thoracic surgery. More recently, a state-of-the-art application of a Deep Learning (DL) technique, namely Deep Reinforcement Learning (DRL) has been proposed by Raghu and co-workers [10] for the definition of continuous-space models for the treatment of sepsis, in a twist that goes beyond the more traditional development-and-use of discriminative classifiers. Support vector machine (SVM) models have also been used for the prediction of sepsis. Kim and co-workers [11] applied them to study the occurrence of sepsis in post-operative patients. Wang *et al.* [12] went a step further to build a DSS for the diagnosis of sepsis.

ML methods have also been used with varying success for the more specific problem of the prediction of mortality caused by sepsis. A diagnostic system for septic shock based on ANNs (Radial Basis Functions -RBF- and supervised Growing Neural Gas) was presented in [14]. Also in this area, Brause and colleagues [15] applied an evolutionary algorithm to an RBF network (the MEDAN Project) to obtain, over a retrospective dataset, a set of predictive attributes for assessing mortality for Abdominal Sepsis. Relevance Vector Machines (RVM), SVM variants with embedded feature selection, were used in [19], while Bayesian Networks were used in [16] and [17]. Finally, kernel methods were used in [18].

The work reported in this paper attempts to identify clinical traits that can be used as predictors of mortality in patients with cardiogenic or septic shock. Missing data imputation techniques and ML models are used, followed by different approaches to FS from different data subsets. Selected features are later used to build causal Bayesian Networks (CBN), in order to reveal potential causal relationships between features and ICU outcome.

The remaining of the paper is structured as follows. Section Materials and methods describes the *ShockOmics* dataset, as well as the FS and structure learning methods used in the reported experiments. Section Experimental Results presents the results for the FS experiments (obtained feature subsets and their classification performances) and causal discovery experiments (CBN structures). The experimental results are discussed in Section Discussion. The last section of the paper draws some conclusions about the investigated clinical problem.

Materials and methods

The ML pipeline presented in this paper is divided into two main experimental phases: FS and causal discovery.

The FS experiments deal with the problem of finding the most promising attributes (clinical traits) for the prediction of mortality related to shock. To do this, three subsets of the *ShockOmics* data were analyzed: *T1*, *T1+T2* and *Full* (description in the following sections). The goal of the experiments was to analyze different feature sets resulting from the FS procedure, and to measure their performance using a Gaussian Naive Bayes model (G-NB). The decision of starting the analyses with four different data subsets of the *ShockOmics* database was made due to the importance of obtaining a stable subset of features that behave consistently at different time-steps and that are clinically relevant. For this reason, each of the datasets targeted different features: some of them because of their low amount of missing values, some of them because they were important at different times. There is an assumption behind the experiments, namely that the feature sets with the best performances are also the most promising for mortality prediction.

The evaluated feature sets were obtained by applying five different FS techniques to the aforementioned datasets: Univariate FS based on ANOVA F-value (*UFS*); recursive feature elimination with Support Vector Classification (*RFE*); recursive feature elimination with cross-validation and Support Vector Classification (*RFECV*), two stage FS (*UFS+RFE*) and Random Forest (*RF*) FS. Missing values were imputed using RF imputation. The identification of the promising feature sets was based on both performance and stability scores.

Finding a subset of attributes that are predictive of mortality due to shock may not suffice in a clinical setting. The causal discovery experiments reported next allow us to go one step further and analyze the causal relationships between the data features that were selected in the previous phase of the ML pipeline. In order to reveal these causal relationships, the Fast Greedy Search algorithm was used. The algorithm was applied twice for each of the feature subsets obtained in the FS experiments. First, the algorithm was applied only to the features, excluding the outcome, and, then, it was applied to both the features and the outcome. The expectation was to find few or no differences between features that did not involve the outcome. That would mean that the CBNs are stable enough to draw some clear conclusions.

For the implementation and more detailed information about the methods, please refer to the corresponding sections.

Dataset description

This study is part of the prospective observational trial *ShockOmics* European research project ("Multiscale approach to the identification of molecular biomarkers in acute heart failure induced by shock", Nr. 602706 - ClinicalTrials.gov Identifier NCT02141607) [34] and was approved by: the Geneva Regional Research Ethics Committee (study number 14-041), the Ethical Committee of Hôpital Erasme-Université Libre de Bruxelles (study number P2014/171), and the Mutua de Terrassa Hospital Institutional Review Board (study number EO/1407). Patients admitted with both cardiogenic and septic shock to the intensive care units (ICU) of Geneva University Hospitals, Geneva, Switzerland (38-bed mixed) and Erasme Univesity Hospital, Brussels, Belgium (36-bed mixed) and Mutua de Terrassa (12-bed mixed) between October 2014 and March 2016 were screened for inclusion criteria. This study included patients diagnosed with septic or cardiogenic shock who were older than 18 years with a Sequential Organ Failure Assessment (SOFA) greater than 2, arterial lactate greater or equal than 2 mmol/l and with a documented source of infection at admission for patients with septic shock. Patients with a high risk of death within the first 24h after admission, systemic immunosuppresion, haematological diseases, metastasic cancer, pre-existing dialysis, decompensated cirrhosis or who had received more than 4 units of red blood cells or any fresh frozen plasma before screening were excluded from the study [34].

The complete database integrates blood samples and hemodynamic recordings from septic shock and cardiogenic shock patients, and from sepsis patients, obtained in ICU [34]. Clinical data and blood samples were collected for analysis at: 1) T1 $<$ 16 h from T0; 2) T2 = 48 h after T0; 3) T3 = day 7 or before discharge or before discontinuation of therapy in case of fatal outcome [35]. Since a lot of patients either died or were discharged during the study, the dataset has a lot of missing values at later time-steps (especially at T3).

The *ShockOmics* dataset (data are available from the ShockOmics Institutional Data Access for researchers who meet the criteria for access to confidential data) has 330 features and 75 observations. 32.1% of all values are missing, and the highest percentage of missing values in one feature is 97%. Fig 1 shows a graphical depiction

Table 1. Clinical data summary.

Parameter	Cardiogenic shock	Septic shock	p-values
Number of patients	25	50	-
Gender (Males:Females)	20:5	34:16	-
Outcome (Alive:Dead)	19:6	38:12	-
Body mass index	26.8 ± 5.93	25.22 ± 5.54	0.053
Total days in ICU	6.60 ± 5.42	7.56 ± 6.54	0.2598
SOFA T1	10.72 ± 4.27	11.50 ± 3.7	0.169
SOFA T2	8.33 ± 4.92	8.77 ± 4.30	0.5015
SOFA T3	5.87 ± 3.68	6.75 ± 4.7	0.1421
APACHE II T1	24.12 ± 9.79	23.16 ± 7.52	0.4378
Lactate levels T1	6.04 ± 4.3	4.84 ± 2.58	0.0179
Lactate levels T2	1.52 ± 0.91	2.1 ± 1.52	0.0013
Lactate levels T3	1.25 ± 0.5	1.43 ± 0.58	0.0197

Summaries of patients populations with cardiogenic and septic shock. Numerical parameters of both populations ($mean \pm std$) are compared against each other with the Welch's t-test, where the null hypothesis is that both parameters have identical mean values.

map of missing values for each observation. The dataset has both numerical and categorical values, but the majority of data is numerical.

Fig 1. Graphical depiction map of missing values in *ShockOmics* dataset. The rows and the columns of the map correspond to observations and features, respectively. The black color represents missing values, the white color corresponds to present values in the *ShockOmics* dataset.

The *ShockOmics* dataset was split into three datasets. The first dataset was named *Full* and it was obtained by filtering features from the original *ShockOmics* dataset. Certain features that did not make sense from the classification viewpoint were manually removed. They were either very general comments in natural language or explicitely revealed some information about the outcome of the patient. Both cases were deemed not to be reliable as features to feed the ML model. In detail, the following features were manually removed: *Reason for admission*; *ICU admission*; *RV area/LV area* (T1, T2, T3); *Microorganisms* (three columns with the same name); *ID*; *Death due to withdrawal of care*; *Mortality 28 days, 100 days* (two columns); *Hospital results*; *Total days in ICU, in Hospital* (two columns). The resulting dataset had 316 features with one target feature.

A research hypothesis is that the closer in time the feature to the final outcome, the better the prediction of mortality. To test this hypothesis, two further datasets were created, namely **T1+T2** and **T1**. Both were obtained from the *Full* dataset by filtering features at certain time-steps. From the *T1+T2* dataset all T3 features were removed,

resulting in 180 features with a target feature. From the *T1* dataset both T2 and T3 time-step features were left out, so the resulting dataset included 104 features and a target feature.

A fourth feature set was built with features that are commonly assumed to be associated with the mortality of patients with shock, according to current practice. This feature set is referred to as *initial feature set* (*IFS*). Table 2 shows features that were included in the IFS set.

Table 2. Initial feature set description.

Name of the feature	Observations	Type
Which type of shock	75	categorical (4 val)
Lactate levels (mmol/L)$_{T1}$	**72**	numerical (cont)
Lactate levels (mmol/L)$_{T2}$	**64**	numerical (cont)
Lactate levels (mmol/L)$_{T3}$	**38**	numerical (cont)
Mean arterial pressure (mmHg)$_{T1}$	75	numerical (cont)
Mean arterial pressure (mmHg)$_{T2}$	**71**	numerical (cont)
Mean arterial pressure (mmHg)$_{T3}$	**44**	numerical (cont)
SOFA$_{T1}$	75	numerical (cont)
SOFA$_{T2}$	**71**	numerical (cont)
SOFA$_{T3}$	**43**	numerical (cont)
APACHE II$_{T1}$	75	numerical (cont)
Result in ICU	75	categorical (2 val)

The columns of the table correspond to the name of the feature, the number of available observations and the type of the feature. *Initial feature set* (IFS) consists of 11 features and the target (*Result in ICU*). Features with missing observations are highlighted in bold in the *Observations* column.

Feature selection

In order to identify promising feature sets, five FS techniques were applied to the four data sets described in the previous section. The resulting feature subsets were used to train the G-NB model and five performance measures were recorded. The first four measures are defined through the numbers of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) :

- **Accuracy**: $\frac{TP+TN}{TP+TN+FP+FN}$;
- **Matthews correlation coefficient *(MCC)***: $\frac{TP\times TN-FP\times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$;
- **Specificity**: $\frac{TN}{(TN+FP)}$;
- **Sensitivity**: $\frac{TP}{(TP+FN)}$;
- **Area Under the Receiver Operating Characteristic Curve *(AUC)***. The average AUC for both labels, weighted by the number of true instances for both labels.

Since this work uses Bayesian Networks to reveal causal relationships between attributes, the G-NB model was chosen as a classifier. Technically, it can be considered as a special case of Bayesian Networks and, unfortunately, conducting all the

experiments with Bayesian Networks was unfeasible because independence conditions remained unknown. The G-NB model was therefore used as the substitute.

The model training stage was performed 100 times using 75% of the data, which were randomly split, with the rest of the data being left for testing. The mean and the standard deviation of the 100 executions (using the test set) were calculated for all five performance measures. Missing values were imputed using RF imputation [43]. The G-NB model [42] implementation used for the experiments was taken from the scikit-learn library (full name of the method:), while the RF imputation was performed in *R* (v3.4.3) [41] (It can be used to impute continuous and/or categorical data including complex interactions and nonlinear relations. *missForest* package).

The scikit-learn [36] library (v0.19.1) in Python (v2.7) was used for the implementation of the FS methods. The default parameters of the methods were used in the experiments, if not otherwise specified. The five FS techniques used in the identification of promising attributes are described next:

- **Univariate feature selection based on ANOVA F-values (*UFS*).** Univariate feature selection [37] works by selecting the best features based on univariate statistical tests. In this case, ANOVA F-tests are used. The UFS was implemented using the sklearn.feature_selection.SelectKBest method that removes all but the k highest scoring features and sklearn.feature_selection.f_classif for the ANOVA F-tests;

- **Recursive feature elimination with SVC (*RFE*).** Given an external estimator that assigns weights to features (e.g., the coefficients of a linear model), the goal of the RFE is to select features by recursively considering smaller and smaller sets of features [38]. The RFE was implemented through the sklearn.feature_selection.RFE method. For the external estimator, the SVC model [39] was used with the help of the sklearn.svm.SVC method;

- **Recursive feature elimination with cross-validation and SVC (*RFECV*).** Feature ranking with RFE and cross-validated selection of the best number of features. Different scores were used for the cross-validation: AUC, recall, accuracy, precision, average precision, *neg log loss*, f1, f1 *micro*, f1 *macro*. The sklearn.-feature_selection.RFECV method allowed the use this FS technique. Once again, the SVC model was used as the external estimator (see previous entry);

- **Two stage FS (*UFS + RFE*).** A combination of two FS techniques from this list. In the first stage, UFS preselects the 80 best features based on the ANOVA F-value and then RFE (SVC) choses the best features in the second stage.

- **Random Forest (*RF*) FS.** This method [38] is based on the trained RF model [40] and its feature rankings. In the experiments, FS was based on 500 runs of the RF model with 10 trees. This procedure was implemented because the RF FS was the only stochastic technique in the list, and multiple experiments were required to achieve a high level of certainty in the results. The final FS was based on the stability scores from all the executions. The most frequent features got in the final feature subset and all the performance evaluation was done using this subset. The sklearn.feature_selection.SelectFromModel method was used to retrieve the feature rankings and the sklearn.ensemble.RandomForestClassifier method to train the RF model.

Once all the feature sets were tested and their performances measured, the stability scores were calculated. For each of the datasets, there are two groups of stability scores.

The first group contains only features obtained with the RF FS. The second group includes the rest of the methods. The difference between these two groups is, again, that the RF FS method is stochastic. It was therefore executed 500 times, each time yielding slightly different feature sets. The methods in the other group were applied only once, since the result is always the same. By consodering these two groups of methods separately, and by giving their stability scores equal weight in the decision making, the difference between methods was compensated.

In each group, stability scores were calculated as frequencies. The number of individual feature occurrences in the selection was divided by the number of FS procedure executions in the group. So, in the RF group, it was divided by 500 and, in the second group, it was divided by 3 plus the number of the RFECV feature sets. The number of the RFECV feature sets varied depending on the dataset. This is because different scores may lead to different feature sets and only the number of unique feature sets was used for calculating the stability scores. In the end, each dataset had two groups of stability scores and each group was represented using the 30 features with highest stability scores.

Structure learning

The Fast Greedy Search (FGES) algorithm was used to obtain the CBN structures that would reveal the interactions between the features. This is an optimized and parallelized version of the Greedy Equivalence Search algorithm. It heuristically searches the space of CBNs and returns the model with the highest value for the Bayesian Information Criterion (BIC) [45] it finds. It starts with an empty graph and adds new edges to increase the Bayesian score. The process stops when there are no edges that increase the score. It then starts removing edges until no single edge removal can increase the score.

The algorithm works on the assumption that the causal process generating the data is accurately modeled by a CBN. Each node in this CBN is a a linear function of its parents, plus a finite additive Gaussian noise term. Each observation in the data is assumed to be independent and obtained by randomly sampling all the variables from the joint distribution. Given all these assumptions, the FGES procedure outputs the CBN structure that contains:

1. an arc X $\rightarrow$ Y, if and only if X causes Y;
2. an edge (-), if and only if either X causes Y or Y causes X;
3. no edge between X and Y, if and only if X and Y have no direct causal relationship between them.

The *Tetrad* software (v6.4.0) was used in the implementation of FGES. The algorithm was applied twice for each of the feature sets obtained during the feature selection phase. First, the algorithm was applied only to the features, excluding the outcome, while the second time it was instead applied to both features and the outcome. The expectation was to have few or no differences between features that did not involve the outcome. That would mean that the CBNs are stable enough to draw clear conclusions. The algorithm used the default parameters and seven maximum discrete categories.

Experimental Results

Feature selection experiments

The FS experiments are divided into three groups. Each group corresponds to its own dataset: *Full*, *T1+T2* and *T1*. All five FS techniques were applied to each dataset. The

results were used to train the *G-NB* model, and the performance of each feature set with the model was evaluated. In this section, the best models from all FS experiments and their accuracy, MCC and AUC are presented. For the complete experiments, results and additional measures such as sensitivity and specificity, please refer to the individual sections of Appendix A. They include full FS results and stability scores for each dataset.

Table 3 shows the comparison of different feature sets: their size, the dataset and the FS method used. The feature sets from the same dataset are grouped together. The additional IFS dataset is used as a baseline to compare the rest of feature sets performances. A naive *majority classifier* that always predicts the most frequent label in the training set, was also added for baseline comparison. This table presents only the best results for each FS method.

Table 3. The performance comparison of the feature sets.

F	Data	Method	Accuracy	MCC	Sensitivity	Specificity	AUC
-	-	Majority class	0.737 ± 0.0 (0.0)	0.0 ± 0.0 (0.0)	0.0 ± 0.0 (0.0)	1.0 ± 0.0 (0.0)	0.5 ± 0.0 (0.0)
11	IFS	-	0.825 ± 0.071 (0.0012)	0.568 ± 0.181 (0.0)	0.697 ± 0.196 (0.0)	0.87 ± 0.082 (1.0)	0.784 ± 0.098 (0.0)
28	T1	UFS	**0.867 ± 0.07** (0.5812)	0.665 ± 0.189 (0.5448)	0.736 ± 0.198 (0.0001)	0.914 ± 0.071 (0.0003)	0.825 ± 0.102 (0.0508)
26	T1	RFE	0.762 ± 0.062 (0.0)	0.223 ± 0.274 (0.0)	0.252 ± 0.311 (0.0)	0.944 ± 0.088 (0.0)	0.598 ± 0.13 (0.0)
31	T1	RFECV	0.754 ± 0.053 (0.0)	0.183 ± 0.25 (0.0)	0.196 ± 0.287 (0.0)	0.954 ± 0.083 (0.0)	0.575 ± 0.117 (0.0)
29	T1	UFS+RFE	0.791 ± 0.077 (0.0)	0.431 ± 0.215 (0.0)	0.406 ± 0.241 (0.0)	0.928 ± 0.126 (0.0003)	0.667 ± 0.099 (0.0)
4	T1	RF	0.846 ± 0.071 (0.1712)	0.615 ± 0.184 (0.012)	0.713 ± 0.194 (0.0)	0.893 ± 0.079 (0.0626)	0.803 ± 0.099 (0.0004)
7	T1	RF	0.834 ± 0.074 (0.0161)	0.608 ± 0.178 (0.0048)	0.763 ± 0.185 (0.0032)	0.86 ± 0.089 (0.4407)	0.811 ± 0.096 (0.0023)
26	T1+T2	UFS	**0.867 ± 0.076** (0.5945)	0.667 ± 0.193 (0.6002)	0.738 ± 0.187 (0.0001)	0.914 ± 0.077 (0.0004)	0.826 ± 0.102 (0.0599)
3	T1+T2	RFE	0.841 ± 0.068 (0.0639)	0.603 ± 0.179 (0.0027)	0.698 ± 0.195 (0.0)	0.891 ± 0.082 (0.0939)	0.795 ± 0.096 (0.0)
11	T1+T2	RFE	0.812 ± 0.079 (0.0)	0.586 ± 0.179 (0.0003)	0.816 ± 0.169 (0.4267)	0.811 ± 0.087 (0.0)	0.813 ± 0.096 (0.0038)
3	T1+T2	RFECV	0.838 ± 0.07 (0.0354)	0.588 ± 0.183 (0.0004)	0.668 ± 0.19 (0.0)	0.899 ± 0.078 (0.0186)	0.783 ± 0.097 (0.0)
28	T1+T2	UFS+RFE	0.785 ± 0.043 (0.0)	0.342 ± 0.211 (0.0)	0.262 ± 0.178 (0.0)	0.972 ± 0.043 (0.0)	0.617 ± 0.081 (0.0)
5	T1+T2	RF	0.854 ± 0.075 (0.5322)	0.646 ± 0.183 (0.179)	0.768 ± 0.169 (0.0039)	0.884 ± 0.083 (0.2656)	0.826 ± 0.094 (0.0495)
9	T1+T2	RF	0.864 ± 0.064 (0.775)	0.658 ± 0.166 (0.3545)	0.718 ± 0.188 (0.0)	0.916 ± 0.075 (0.0002)	0.817 ± 0.092 (0.0078)
7	Full	UFS	0.861 ± 0.083 (1.0)	0.681 ± 0.184 (1.0)	0.834 ± 0.15 (1.0)	0.87 ± 0.094 (1.0)	0.852 ± 0.092 (1.0)
19	Full	UFS	0.869 ± 0.068 (0.4568)	0.681 ± 0.171 (1.0)	0.796 ± 0.167 (0.0921)	0.895 ± 0.072 (0.0361)	0.845 ± 0.09 (0.5871)
23	Full	RFE	0.785 ± 0.077 (0.0)	0.431 ± 0.212 (0.0)	0.502 ± 0.187 (0.0)	0.886 ± 0.088 (0.2155)	0.694 ± 0.099 (0.0)
19	Full	RFECV	0.743 ± 0.079 (0.0)	0.283 ± 0.227 (0.0)	0.356 ± 0.187 (0.0)	0.881 ± 0.095 (0.4115)	0.619 ± 0.098 (0.0)
15	Full	UFS+RFE	**0.873 ± 0.068 (0.2648)**	0.686 ± 0.174 (0.8437)	0.764 ± 0.188 (0.004)	0.912 ± 0.075 (0.0006)	0.838 ± 0.096 (0.2937)
14	Full	RF	0.87 ± 0.072 (0.4137)	0.683 ± 0.169 (0.9363)	0.77 ± 0.175 (0.006)	0.906 ± 0.084 (0.0048)	0.838 ± 0.092 (0.2832)

Best feature sets obtained in the FS experiments and their performance. The columns correspond to the number of features, the data that was used to obtain the feature set, the FS method and five performance measures ($mean \pm std(p - value)$). Welch's t-test is used to obtain p-values for the null hypothesis that two performance measures have identical values. Each measure was tested against the same measure but of the *Full (7, UFS)* feature set. The *IFS* is used as a baseline for comparison. The best results for accuracy are highlighted in bold.

As can be seen from Table 3, the UFS and the RF models produce consistently good results. Their performance is better than the models with the IFS, and this is especially noticeable when it comes to the MCC. The best results in the experiments were found for the *UFS* FS method in the *Full* dataset, while the second best was the *UFS+RFE* technique with the same dataset. This is interesting, since *UFS+RFE* showed poor performance in other two datasets. The *RFECV* features achieved the worst results. It uses the SVC model for cross-validation, but the performance was measured with the *G-NB*. It was impossible to use the *G-NB* model, since it cannot evaluate the importance of features. When the SVC model was used for testing the performance, it showed slightly better results. The *RFE* feature showed reasonably good performance but worse than the *UFS* and the *RF*.

In order to obtain the most promising feature sets, the results from both tables were filtered based on their performance (see Table 4). For this purpose, different performance thresholds were used. Only the feature sets with accuracy ≥ 0.86, MCC ≥ 0.68 and AUC ≥ 0.84 were selected for the next experiment. All filtered feature sets and their individual features can be found in the list below.

Table 4. The performance of the filtered feature sets.

F	Data	Method	Accuracy	MCC	Sensitivity	Specificity	AUC
28	T1	UFS	0.867 ± 0.07 (0.5812)	0.665 ± 0.189 (0.5448)	0.736 ± 0.198 (0.0001)	**0.914 ± 0.071 (0.0003)**	0.825 ± 0.102 (0.0508)
26	T1+T2	UFS	0.867 ± 0.076 (0.5945)	0.667 ± 0.193 (0.6002)	0.738 ± 0.187 (0.0001)	**0.914 ± 0.077 (0.0004)**	0.826 ± 0.102 (0.0599)
5	T1+T2	RF	0.854 ± 0.075 (0.5322)	0.646 ± 0.183 (0.179)	0.768 ± 0.169 (0.0039)	0.884 ± 0.083 (0.2656)	0.826 ± 0.094 (0.0495)
7	Full	UFS	0.861 ± 0.083 (1.0)	0.681 ± 0.184 (1.0)	**0.834 ± 0.15 (1.0)**	0.87 ± 0.094 (1.0)	**0.852 ± 0.092 (1.0)**
19	Full	UFS	0.869 ± 0.068 (0.4568)	0.681 ± 0.171 (1.0)	0.796 ± 0.167 (0.0921)	0.895 ± 0.072 (0.0361)	0.845 ± 0.09 (0.5871)
15	Full	UFS+RFE	**0.873 ± 0.068 (0.2648)**	**0.686 ± 0.174 (0.8437)**	0.764 ± 0.188 (0.004)	0.912 ± 0.075 (0.0006)	0.838 ± 0.096 (0.2937)
14	Full	RF	0.87 ± 0.072 (0.4137)	0.683 ± 0.169 (0.9363)	0.77 ± 0.175 (0.006)	0.906 ± 0.084 (0.0048)	0.838 ± 0.092 (0.2832)

The columns of the table correspond to the number of features, the data that was used to obtain the feature set, the FS method used and three performance measures: the accuracy, the MCC and the AUC. The filtering was based on the following criteria: the accuracy ≥ 0.85, the MCC ≥ 0.665, the AUC ≥ 0.825. The best results in the performance measures columns are highlighted in bold.

```
T1 (28, UFS): APACHE II_T1, SOFA_T1, Respiratory rate_T1, X
    Norepinephrine mg kg min _T1, Number of affected organs, Tidal
    volume VT _T1, LA dilatation by eyeballing_T1=No, LA dilatation by
    eyeballing_T1=Yes, Glasgow Coma Scale_T1, Fluid Balance ml _T1,
    Platelet count_T1, LV Dilatation LVEDV _T1=Moderate, E A ratio_T1=1
    39, Aortic Valve Stenosis_T1=Moderate, Base Excess mmol L _T1, Urine
     Output mL day _T1, Anuria and or RRT_T1=Yes, Anuria and or RRT_T1=
    No, LVOT average diameter mm _T1, Mode_T1=Volume controlled, Lactate
     levels mmol L _T1, Is the patient under other drugs_T1=Yes, Is the
    patient under other drugs_T1=No, Hypotension,MAP or SBP decrease _T1
    =Yes, Site of sampling=Urine, Hypotension,MAP or SBP decrease _T1=No
    , Mode_T1=Pressure controlled, Sedation Scale SAS _T1;

T1+T2 (26, UFS): APACHE II_T1, Anuria and or RRT_T1=Yes, Aortic Valve
```

```
    Stenosis_T1=Moderate, Base Excess mmol L _T1, Creatinine mg dL _T2,
    E A ratio_T1=1.39, FiO2_T2, Fluid Balance ml _T1, Glasgow Coma
    Scale_T1, Glasgow Coma Scale_T2, K mmol L _T2, LA dilatation by
    eyeballing_T1=No, LA dilatation by eyeballing_T1=Yes, LV Dilatation
    LVEDV _T1=Moderate, Lactate levels mmol L _T2, Number of affected
    organs, Platelet count_T1, Platelet count_T2, RBC count_T2,
    Respiratory rate_T1, SOFA_T1, SOFA_T2, Sat O2 FiO2_T2, Sedation
    Scale SAS _T2, Tidal volume VT _T1, Tracheal Intubation_T2=No;

T1+T2 (5, RF): SOFA_T2, Lactate levels mmol L _T2, APACHE II_T1,
    Respiratory rate_T1, Urine Output mL day _T1;

Full (7 and 19, UFS): APACHE II_T1, Creatinine mg dL _T3, E e _T3,
    FiO2_T2, FiO2_T3, Glasgow Coma Scale_T2, Heart rate bpm _T3 1,
    Hypotension SBP 90 mmHg or MAP 70 mmHg or SBP decrease 40 mmHg _T2=
    No, Hypotension SBP 90 mmHg or MAP 70 mmHg or SBP decrease 40 mmHg
    _T2=Yes, Inferior vena cava distensibility index _T3, LA dilatation
    by eyeballing_T1=No, LA dilatation by eyeballing_T1=Yes, Lactate
    levels mmol L _T2, Na Ur_T3, Number of affected organs, PCT_Value mg
     mL _T2, RBC count_T3, Respiratory rate_T1, SOFA_T1;

Full (15, UFS+RFE): Inferior vena cava distensibility index _T3, LA
    dilatation by eyeballing_T1=No, LA dilatation by eyeballing_T1=Yes,
    Is the patient under other drugs_T1=No, Glasgow Coma Scale_T3, Is
    the patient under other drugs_T1=Yes, Number of affected organs,
    SOFA_T3, Pulmonary artery systolic pressure TR jet by CW CVP mmHg
    _T2, Glasgow Coma Scale_T1, Lactate levels mmol L _T1, SOFA_T2,
    PCT_Value mg mL _T3, Hematocrit _T3, APACHE II_T1;

Full (14, RF): Inferior vena cava distensibility index _T3, Lactate
    levels mmol L _T2, SOFA_T2, SOFA_T3, Na Ur_T3, APACHE II_T1, FiO2_T3
    , PCT_Value mg mL _T1, Tricuspid regurgitation maximal velocity by
    CW cm s _T3, X Norepinephrine mg kg min _T2, Respiratory rate_T1,
    Urine Output mL day _T1, Sat O2 FiO2_T3, SOFA_T1;
```

List of promising feature sets: the subset of data that was used to obtain the feature set; the size and the applied FS technique are in the parenthesis.

Causal discovery

For each of the feature sets there are two corresponding CBNs. The first one (the CBN "a'are was built without betweene the target feature (*is an Result ion of the ICU*), the second one (the CBN "b") included all features and the target.

Fig 2 shows the CBNs for the *T1 (28, UFS)* feature set. This is the biggest feature set in the experiments; however, eight and seven nodes are detached from the rest of the network in the first and the second CBN respectively. Adding the target feature makes *LA dilatation by eyeballing* at T1 connected to *LVOT average diameter* at T1 with undirected and to *X Norepinephrine* at T1 with directed edge. There are no other differences between the two CBNs aside the outcome, which is an indication of the high stability of the structures. *Result in ICU* is connected to *APACHE II* at T1 and *X Norepinephrine* at T1.

Fig 3 shows the CBNs for the *T1+T2 (26, UFS)* feature set. Even though the feature set is almost as big as in the previous case, it only has four disconnected nodes

Fig 2. The *T1 (28, UFS)* CBN. The CBN for the *T1 (28, UFS)* feature set obtained: a) without the target feature; 2) with the target feature.

in both CBNs now. Adding the outcome results in removing the directed edge from *Tracheal intubation* at T2 to *FiO2* at T2. Beside this, there are no other differences between CBNs. The target is influenced by *SOFA* at T2 and at the same time *Result in ICU* causes *FiO2* at T2 and *Sedation Scale SAS* at T2. Most of the edges in both CBNs have established causal directions.

Fig 4 shows the CBNs for the *T1+T2 (5, RF)* feature set. This is the smallest feature set in the experiments, and previously showed good performance. There are no established causal directions in the CBNs, and adding the target does not induce any change. The only connection that *Result in ICU* has is the one with *SOFA* at T2.

Figure reffg:cbnf7 shows the CBNs for the *Full (7, UFS)* feature set. Once again, the CBNs do not have any established directions for the edges. This time both CBNs have nodes that are detached from the network. *Heart rate* at T3 and *E e* at T3 are disconnected in the first and the second CBNs respectively. Adding the outcome removes the edge between *Glasgow Coma Scale* at T2 and *FiO2* at T2 and the connection between *FiO2* at T3 and *E e* at T3. According to the second CBN, *Result in ICU* can be potentially influenced by *APACHE II* at T1, *FiO2* at T3 and *Heart rate* at T3.

Fig 6 shows the CBNs for the *Full (19, UFS)* feature set. The CBNs for this feature set are noticeably less stable compared to the previous structures. Both of the networks have *LA dilatation by eyeballing* at T1 and *Number of affected organs* detached from

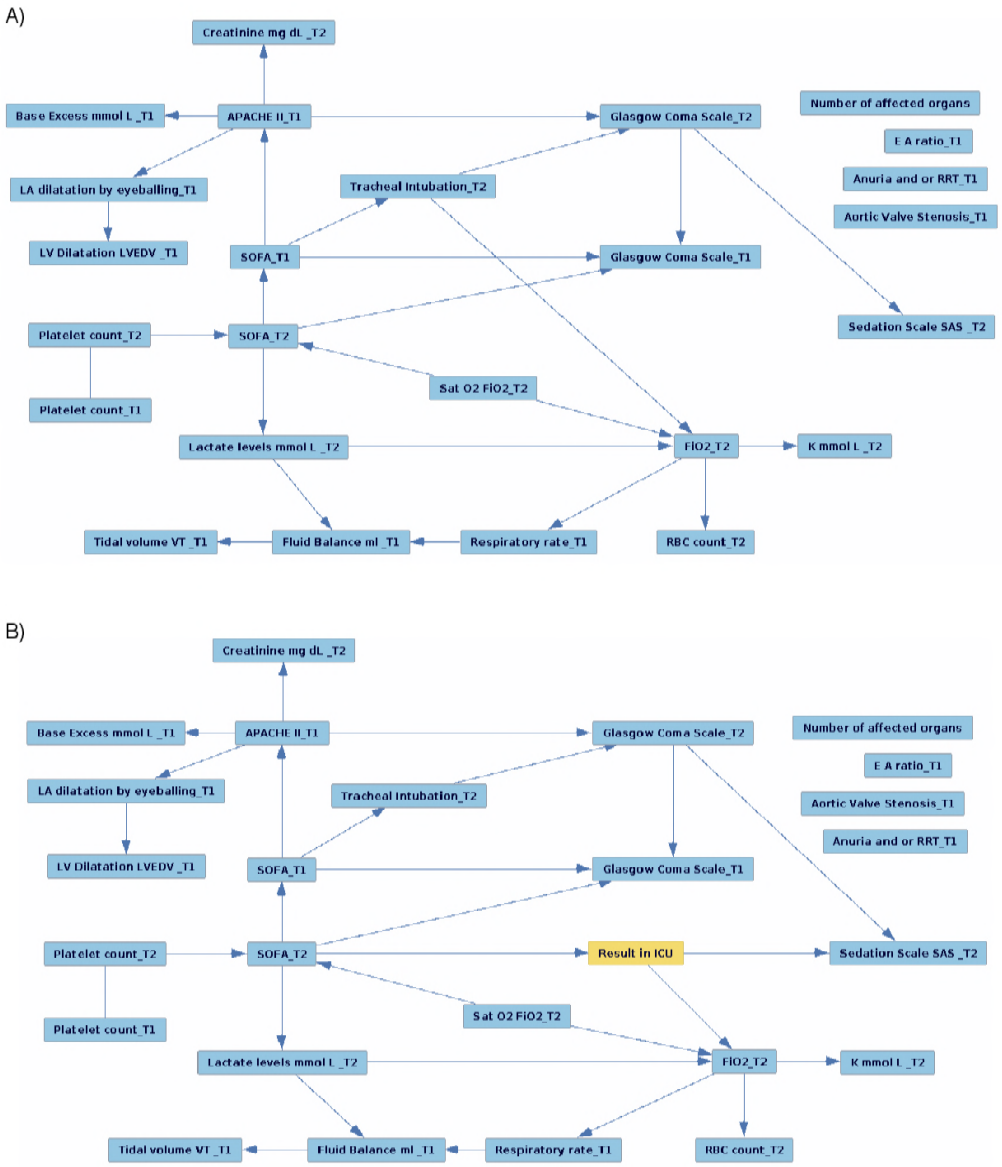

Fig 3. The now *T1+T2 (26, UFS)* CBN. The CBN for the *T1+T2 (26, UFS)* feature set obtained: a) without the target feature,; b) with the target feature.

Fig 4. The *T1+T2 (5, RF)* CBN. The CBN for the *T1+T2 (5, RF)* feature set obtained: a) without the target feature; b) with the target feature.

the rest of the structure. Adding the outcome changes a lot in some parts of the CBN. It completely removes connection between *PCT_Value* at T2 and *Respiratory rate* at T1, between *Hypotension, MAP or SBP decrease* at T2 and *Lactate levels* at T2, between

Fig 5. The *Full (7, UFS)* CBN. The CBN for the *Full (7, UFS)* feature set obtained: a) without the target feature; b) with the target feature.

Inferior vena cava distensibility index at T3 and *SOFA* at T1, and between *Inferior vena cava distensibility index* at T3 and *Creatinine* at T3. Aside of this, causality directions of some edges become clear. This time *Result in ICU* is influenced by *Inferior vena cava distensibility index* at T3 and it causes *FiO2* at T2, *Lactate levels* at T2, *SOFA* at T1 and *Creatinine* at T3.

Fig 7 shows the CBNs for the *Full (15, UFS+RFE)* feature set. The CBNs have very stable connections, and the only not established causal direction is between *Pulmonary artery systolic pressure TR jet by CW* at T2 and Is THE PATIENT UNDER OTHER DRUGS. This connection is also detached from the rest of the network, as well as the *Number of affected organs* node. There are no differences between two CBNs, which means that the structures are stable. *Result in ICU* is only influenced by *Inferior vena cava distensibility index* at T3.

Finally, Fig 8 shows the CBNs for the *Full (14, RF)* feature set. Adding the outcome significantly changes the structures. The connections between *X Norepinephrine* at T2 and *Lactate levels* at T2 are removed, and edges between *Urine Output* at T1 and *PCT_Value* at T1, and between *Na Ur* at T3 and *APACHE II* at T1 are added. In the first CBN all connections are undirected and with the target feature most of them become clear. *Result in ICU* is connected to *FiO2* at T3 and *SOFA* at T1, and *Tricuspid regurgitation maximal velocity by CW* at T3 and *Inferior vena cava distensibility index* at T3 are the only potential influences.

In most of CBNs pairs, the presence of the target value did not generate much change. In some cases, new edges appeared, but, most of the time, only a few edges changed their directions. It should be noted that the majority of edges were undirected. This might be because of the small amount of data available for model generation. Nevertheless, some clear patterns are noticeable throughout the experiments.

Discussion

Analyzing the results of the FS experiments, one may notice that the RF stability scores have a lot of reoccurring features, while the scores for other FS methods are very diverse, instead. The subsequent causal discovery experiments indicate that the RF scores are more accurate in identifying important features for mortality prediction. Features like *SOFA*, *APACHE II*, *Lactate levels*, *Respiratory rate*, *Urine Output* and *X Norepinephrine* usually scored high. In most of the cases, *SOFA* and *APACHE II* were also highly rated by other FS methods. Additionally, it seems that certain features are highly valuable only at certain time steps. For example, *Lactate levels* seems to have the most predictive value at T2, *X Norepinephrine* and *Respiratory rate* are particularly valuable at T1, *Urine Output* – at T1 and T2 and *Heart rate* at T3. These hypotheses were further tested in the causal discovery experiments.

Although such patterns of features are noticeable, the differences between feature

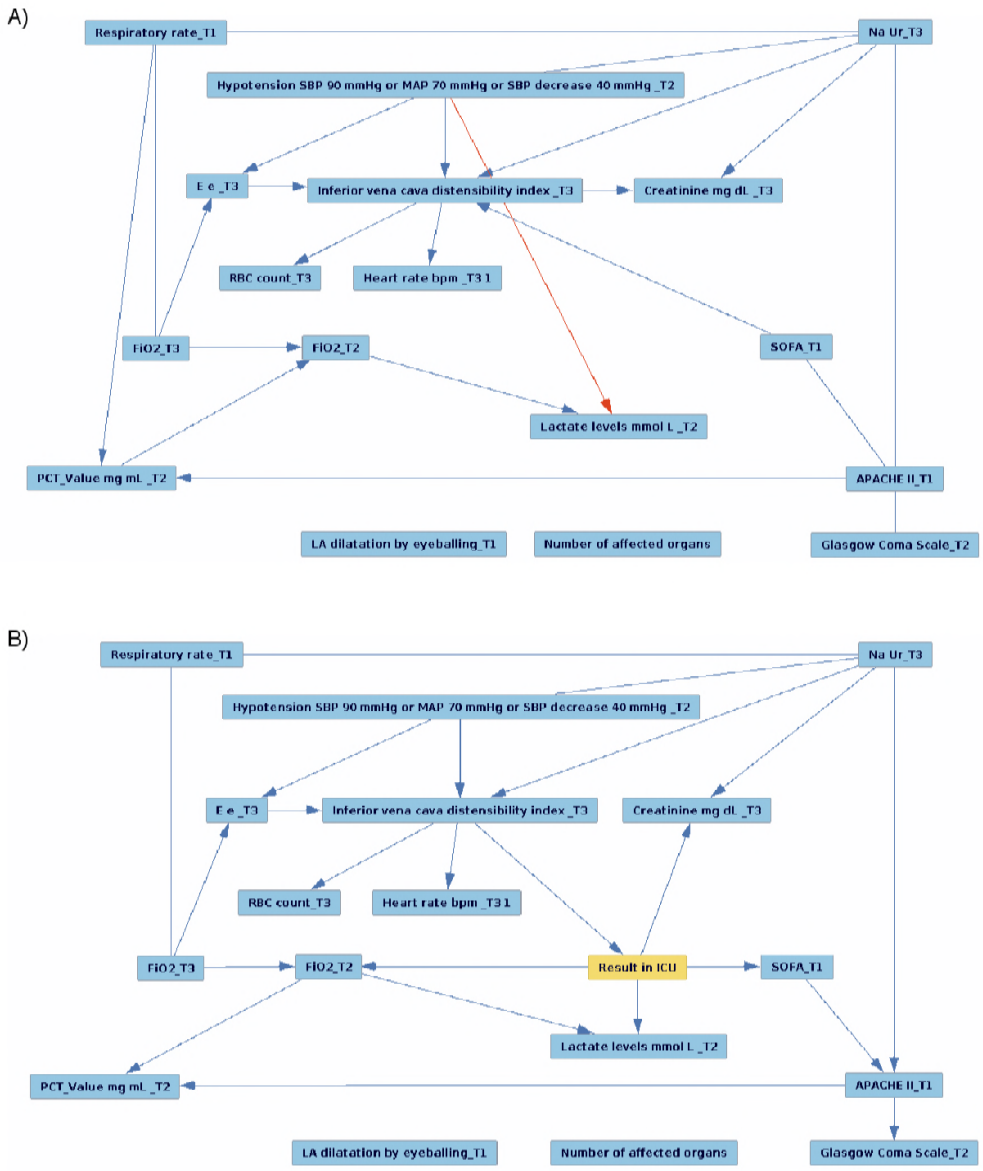

Fig 6. The *Full (19, UFS)* CBN. The CBN for the *Full (19, UFS)* feature set obtained: a) without the target feature; b) with the target feature.

sets in different datasets are also important. In the experiments, the performance measures of *Full (7, UFS)* were tested against the measures of other feature sets, using the Welch's t-test. The null hypothesis was that each pair of performance measures has identical values. Setting the threshold for the p-value to 0.05 shows that there is a significant chance that the best feature sets in the all three datasets perform similarly, at least in terms of accuracy, MCC and AUC. When more time steps are considered, less features are necessary to show such high performance.

The CBN structures revealed connections between the features and the outcome. The assumption behind the evaluation of the most important clinical attributes for outcome prediction was that such features are the closest to the outcome in a graph. The direction of the edge from feature to the outcome is also a good indicator that the

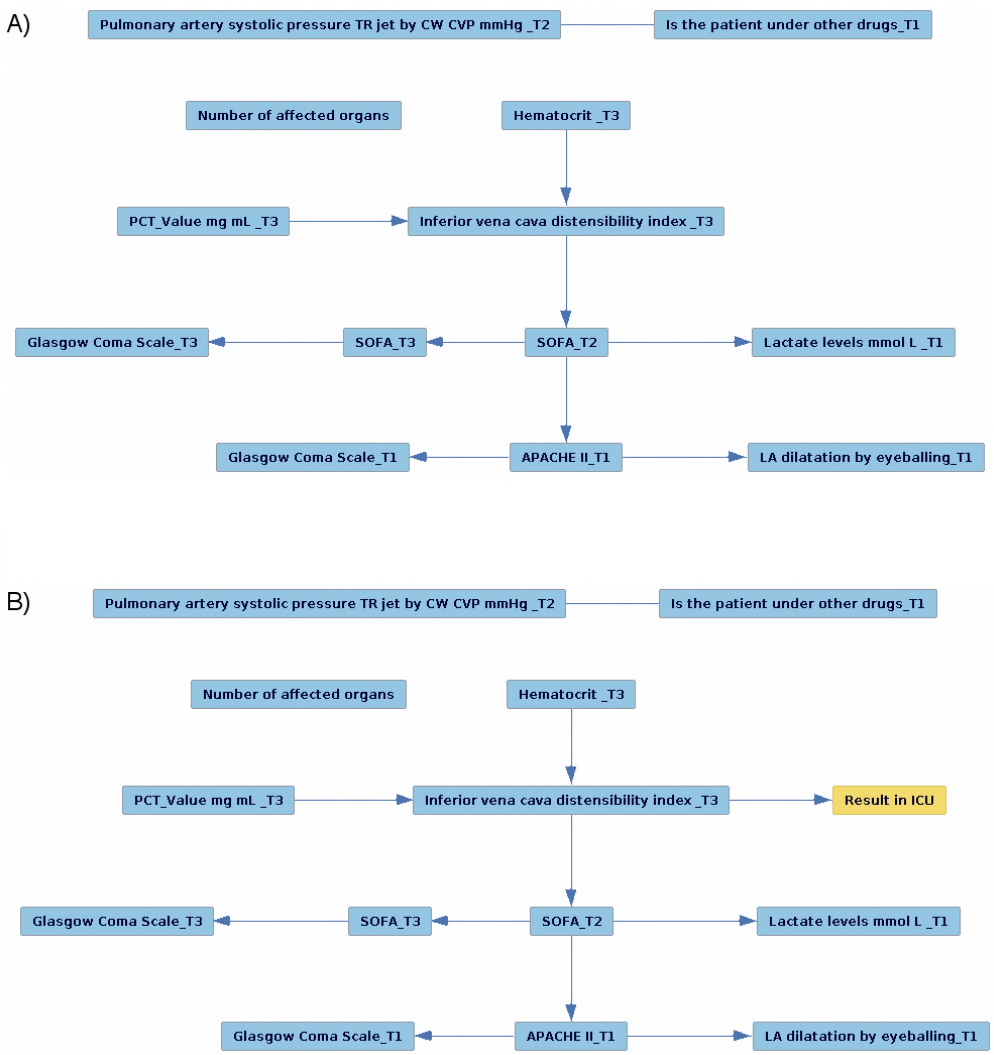

Fig 7. The *Full (15, UFS+RFE)* CBN. The CBN for the *Full (15, UFS+RFE)* feature set obtained: a) without the target feature; b) with the target feature.

feature is important.

According to the structures that contain *Result in ICU*, *Inferior vena cava distensibility index* at T3, followed by *SOFA* at T2 and *APACHE II* at T1 were considered the most important features most of the times. In all CBNs that involved *Inferior vena cava distensibility index* at T3 it either directly influenced or was connected to the outcome. It appeared in all three best feature sets of the *Full* dataset. *SOFA* at T2 appeared connected to the target feature in both *T1+T2* feature sets, being influenced by the outcome in the *T1+T2 (26, UFS)* feature set. *APACHE II* at T1 was the only direct influence on *Result in ICU* in the *T1 (28, UFS)* feature set, and its influence was later confirmed by the undirected edge with the outcome in the *Full (7, UFS)*, that achieved the best performance in the experiments.

Other features that could potentially influence *Result in ICU* are *Heart rate* at T3 and *FiO2* at T3 in the *Full (7, UFS)* feature set, and *Tricuspid regurgiation maximal velocity by CW* at T3 in the *Full (14, RF)* feature set. The rest of the features, that were added to the list of important attributes, were influenced by the outcome. This

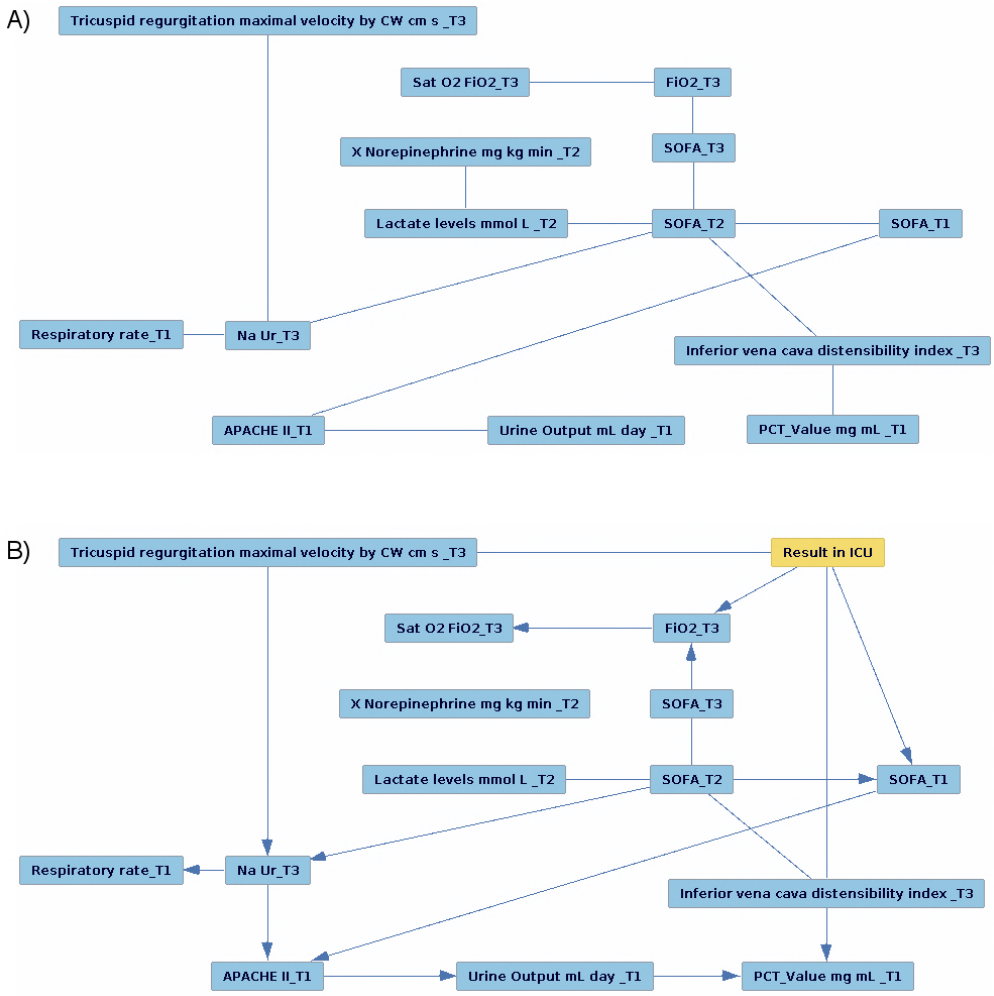

Fig 8. The *Full (14, RF)* CBN. The CBN for the *Full (14, RF)* feature set obtained: a) without the target feature; b) with the target feature.

close connection to the target feature makes them valuable for mortality prediction, as well. Such feature include *X Norepinephrine* at T1, *FiO2* at T2, *Sedation Scale SAS* at T2, *FiO2* at T2, *Lactate levels* at T2, *SOFA* at T1, *Creatinine* at T3. *SOFA* at T1 and *FiO2* at T2 appeared close to *Result in ICU* twice, which increases the confidence in their importance.

These findings agree with other studies that support the use of certain attributes for mortality prediction. For instance, the SOFA score has previously been shown to have a significant prognostic value for in-hospital mortality prediction [46] [47]. The APACHE II score is often used for the outcome prediction too [48–50]. And in the study by Jansen *et al.* [51], targeting a decrease of at least 20% in the blood lactate level over a 2-hour period seemed to be associated with reduced in-hospital mortality in patients with shock. However, some of the attributes that are also related to the mortality in patients with shock, like mean blood pressure [52], mean arterial pressure [53], or low LVEF [54] did not show significant predictive value in this study. The full list of features that were considered valuable for the mortality prediction task can be found below.

```
Features from the Causal Discovery experiments: Inferior vena cava
    distensibility index T3, SOFA T2, APACHE II T1, FiO2 T3, Heart rate
```

```
    T3, Tricuspid regurgitation maximal velocity by CW T3, SOFA T1, FiO2
     T2, Lactate levels T2, X Norepinephrine T1, Sedation Scale SAS T2,
    Creatinine T3.

Features with high stability scores: SOFA T3, Urine Output T1, Urine
    Output T2, Respiratory rate T1, E wave T1.
```

List of attributes that were considered promising for the outcome prediction in patients with shock.

Conclusion

In this paper, we have presented several classifiers for the assessment of the prognosis of septic and cardiogenic shock in the ICU at different time-points. These classifiers are the result of the application of different FS techniques over the available clinical data. For the acute phase, we have obtained a classifier that leverages upon well established attributes for septic shock such as the SOFA or APACHE II scores and treatment with Norepinephrine, which is closely related to the cardiovascular SOFA score. This classifier for the acute phase of shock also resorts to other valuable attributes such as those obtained through echocardiography during cardiogenic shock (e.g. LA dilatation, LV dilatation or aortic valve stenosis). Other important features related to the cardiovascular function such as significant decrease in blood pressure are also significant. All these significant attributes are very closely related to the new oficial definition of shock and its corresponding management guidelines. However, our causal discovery experiments did not yield a strong dependency with the outcome for the echocardiographic data. This may be due to a limitation in the sample size available for the study. The best classifier for the acute phase yielded an accuracy of 0.867, a sensitivity of 0.736 and a specificity of 0.914. The accuracy results were not statistically different from those obtained from the whole dataset.

From our experiments, it is clear that sensitivity improves as patient's history unfolds in the ICU. Once again, for T1+T2 and T3, the presence of echocardiographic attributes as well as other attributes related to blood pressure are also important for the prediction of risk of death.

The main conclusion of this study is that it is possible to predict risk of death in the acute phase of septic and cardiogenic shock with quite acceptable results by taking into consideration the attributes routinely measured through echocardiography in the ICU. Use of this data for assessing the prognosis of patients is considered valuable for the clinical management of patients with shock in the ICU.

These results presented here open up a new line of research for the study of the pathophysiology of shock when combined with other data related to organ dysfunction beyond the commonly used scales such as the SOFA score. These other data may include transcriptomics, proteomics and metabolomics features within the scope of the original Shockomics European research project, which shall be explored in future research.

Acknowledgments

This research was supported by the European FP7 project Shockomics (Nr. 602706), Multiscale approach to the identification of molecular biomarkers in acute heart failure induced by shock, and Spanish research project TIN2016-79576-R.

References

1. Conant GC, Wolfe KH. Turning a hobby into a job: how duplicated genes find new functions. Nat Rev Genet. 2008 Dec;9(12):938–950.
2. Ohno S. Evolution by gene duplication. London: George Alien & Unwin Ltd. Berlin, Heidelberg and New York: Springer-Verlag.; 1970.
3. Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM. Successive increases in the resistance of Drosophila to viral infection through a transposon insertion followed by a Duplication. PLoS Genet. 2011 Oct;7(10):e1002337.
4. Denai M, Mahfouf M, Ross J. A fuzzy decision support system for therapy administration in cardiovascular intensive care patients. In: Proceedings of the FUZZ-IEEE; 2007. p.1-6.
5. Paetz H. Metric rule generation with septic shock patient data. In: Proceedings of the ICDM; 2001. p.637-8.
6. Paetz J. Intersection based generalization rules for the analysis of symbolic septic shock patient data. In: Proceedings of the ICDM; 2002. p. 673-6.
7. Schuh ChJ. Sepsis and septic shock analysis using neural networks. In: Proceedings of the Annual Meeting of the NAFIPS; 2007. p. 650-4.
8. Duhamel A, Beuscart R, Demongeot J, Mouton Y. SES (Septicemia Expert System): knowledge validation from data analysis. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS) vol.3. 1988; p.1400-1.
9. Ross JJ, Mason DG, Paterson IG, Linkens DA, Edwards ND. Development of a knowledge-based simulator for haemodynamic support of septic shock. In: IEEE Colloquium on Simulation in Medicine (Ref. No. 1998/256). 1998; 3/1-3/4.
10. Raghu A, Komorowski M, Celi LA, Szolovits P, Ghassemi M. Continuous State-Space Models for Optimal Sepsis Treatment-a Deep Reinforcement Learning Approach. arXiv preprint. 2017; arXiv:1705.08422.
11. Kim J, Blum J, Scott, C. Temporal features and kernel methods for predicting sepsis in postoperative patients. Technical Report, University of Michigan, USA. 2010.
12. Wang SL, Wu F, Wang BH. Prediction of severe sepsis using SVM model. In: Advances in Experimental Medicine and Biology Series, vol.680(1), Springer; 2010. p. 75–81.
13. Tang CHH, Middleton PM, Savkin AV, Chan GSH, Bishop S, Lovell NH. Non-invasive classification of severe sepsis and systemic inflammatory response syndrome using a nonlinear support vector machine: a preliminary study. Physiol. Meas. 2010;31:775–93.
14. Brause R, Hamker F, Paetz J. Septic shock diagnosis by neural networks and rule based systems. Studies in Fuzziness and Soft Computing, vol.96; 2002. p.323–56.
15. Brause R, Hanisch E, Paetz J, Arlt B. Neural networks for sepsis prediction - the MEDAN project. Journal fur Anasthesie und Intensivbehandlung. 2004;11:40–3.

16. Ribas VJ, Caballero-López J, Sáez de Tejada A, Ruiz-Rodríguez JC, Ruiz-Sanmartín A, Rello J, Vellido A. On the use of graphical models to study ICU outcome prediction in septic patients treated with statins. In: Biganzoli E, Vellido A, Ambrogi F, Tagliaferri R, editors. CIBB 2011, LNCS/LNBI vol.7548; 2012. p.98-111.

17. Morales C, Vellido A, Ribas V. Applying conditional independence maps to improve sepsis prognosis. In: Data Mining in Biomedical Informatics and Healthcare (DMBIH) Workshop. IEEE International Conference on Data Mining (ICDM), 2016.

18. Ribas V, Vellido A, Romero E, Ruiz-Rodríguez JC. Sepsis mortality prediction with quotient basis kernels. Artif. Intell. Med. 2014;61(1):45–52.

19. Ribas V, Ruiz-Rodríguez JC, Wojdel A, Caballero-López J, Ruiz-Sanmartín A, Rello J, Vellido A. Severe sepsis mortality prediction with relevance vector machines. In: Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2011. p.100-3.

20. Dybowski R, Weller P, Chang R, Gant V. Prediction of outcome in critically ill patients using artificial neural network synthesised by genetic algorithm. Lancet 1996;347(9009):1146–50.

21. Wong LS, Young JD. A comparison of ICU mortality prediction using the APACHE II scoring system and artificial neural networks. Anaesthesia. 1999;54(11):1048–54.

22. Clermont G, Angus D, DiRusso S, Griffin M, Linde-Zwirble W. Predicting hospital mortality for patients in the intensive care unit: a comparison of artificial neural networks with logistic regression models. Crit. Care Med. 2001;29(2):291–6.

23. Citi L, Barbieri R. Physionet 2012 Challenge: predicting mortality of ICU patients using a cascaded SVM-GLM paradigm. Comput. Cardiol. 2012;39:257–60.

24. Nachimuthu SK, Wong A, Haug PJ. Modeling glucose homeostasis and insulin dosing in an intensive care unit using dynamic Bayesian networks. In: Proceedings of the AMIA Annual Symposium; 2010. p.532.

25. Nemati S, Adams R. Identifying outcome-discriminative dynamics in multivariate physiological cohort time series. In: Advanced State Space Methods for Neural and Clinical Data. Cambridge, U.K. Cambridge Univ. Press; 2015. p.283.

26. Lasko TA, Denny JC, Levy MA. Computational phenotype discovery using unsupervised feature learning over noisy, sparse, irregular clinical data. PloS ONE 2013;8(6):e66341.

27. Fialho AS, Celi LA, Cismondi F, Vieira SM, Reti SR, Sousa JMC, Finkelstein SN. Disease-based modeling to predict fluid response in intensive care units. Methods Inf. Med. 2013;52(6):494–502.

28. Vincent JL, De Backer D. Circulatory shock. New England Journal of Medicine. 2013;369(18):1726–1734.

29. Sterling SA, Puskarich MA, Glass AF, Guirgis F, Jones AE. The impact of the Sepsis-3 septic shock definition on previously defined septic shock patients. Critical care medicine. 2017;45(9):1436–1442.

30. Nguyen HL, Yarzebski J, Lessard D, Gore JM, McManus DD, Goldberg RJ. Ten-Year (2001–2011) Trends in the Incidence Rates and Short-Term Outcomes of Early Versus Late Onset Cardiogenic Shock After Hospitalization for Acute Myocardial Infarction. Journal of the American Heart Association. 2017;6(6):e005566.

31. McNally R, Mair P, Mugno B, Riemann B. Co-morbid obsessive–compulsive disorder and depression: A Bayesian network approach. Psychological medicine. 2017;47(7):1204–1214.

32. Luo Y, El Naqa I, McShan DL, Ray D, Lohse I, Matuszak MM, et al. Unraveling biophysical interactions of radiation pneumonitis in non-small-cell lung cancer via Bayesian network analysis. Radiotherapy and Oncology. 2017;123(1):85–92.

33. Cooper G, Cai C, Lu X. Tumor-specific Causal Inference (TCI): A Bayesian Method for Identifying Causative Genome Alterations within Individual Tumors. bioRxiv. 2018; p. 225631.

34. Ferrario M, Aletti F, Baselli G. The EU ShockOmics Project International Workshop at ICCAI'17; 2017.

35. Aletti F, Conti C, Ferrario M, Ribas V, Pinto BB, Herpain A, et al. ShockOmics: multiscale approach to the identification of molecular biomarkers in acute heart failure induced by shock. Scandinavian journal of trauma, resuscitation and emergency medicine. 2016;24(1):9.

36. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12:2825–2830.

37. Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. bioinformatics. 2007;23(19):2507–2517.

38. Lin X, Wang Q, Yin P, Tang L, Tan Y, Li H, et al. A method for handling metabonomics data from liquid chromatography/mass spectrometry: combinational use of support vector machine recursive feature elimination, genetic algorithm and random forest for feature selection. Metabolomics. 2011;7(4):549–558.

39. Suykens JA, Vandewalle J. Least squares support vector machine classifiers. Neural processing letters. 1999;9(3):293–300.

40. Liaw A, Wiener M, et al. Classification and regression by randomForest. R news. 2002;2(3):18–22.

41. R Core Team. R: A Language and Environment for Statistical Computing; 2017. Available from: `https://www.R-project.org/`.

42. Chan TF, Golub GH, LeVeque RJ. Updating formulae and a pairwise algorithm for computing sample variances. In: COMPSTAT 1982 5th Symposium held at Toulouse 1982. Springer; 1982. p. 30–41.

43. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics. 2011;28(1):112–118.

44. Stekhoven DJ. missForest: Nonparametric missing value imputation using random forest. Astrophysics Source Code Library. 2015;.

45. Raftery AE. Bayesian model selection in social research. Sociological methodology. 1995; p. 111–163.

46. Rivera-Fernández R, Nap R, Vázquez-Mata G, Miranda DR. Analysis of physiologic alterations in intensive care unit patients and their relationship with mortality. Journal of critical care. 2007;22(2):120–128.

47. Raith EP, Udy AA, Bailey M, McGloughlin S, MacIsaac C, Bellomo R, et al. Prognostic accuracy of the SOFA score, SIRS criteria, and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. Jama. 2017;317(3):290–300.

48. Huang Y, Chen J, Zhong S, Yuan J. Role of APACHE II scoring system in the prediction of severity and outcome of acute intracerebral hemorrhage. International Journal of Neuroscience. 2016;126(11):1020–1024.

49. Agarwal A, Choudhary GS, Bairwa M, Choudhary A. Apache II scoring in predicting surgical outcome in patients of perforation peritonitis. International Surgery Journal. 2017;4(7):2321–2325.

50. Lee H, Lim C, Hong H, Ju J, Jeon Y, Hwang J, et al. Efficacy of the APACHE II score at ICU discharge in predicting post-ICU mortality and ICU readmission in critically ill surgical patients. Anaesthesia & Intensive Care. 2015;43(2).

51. Jansen TC, van Bommel J, Schoonderbeek FJ, Sleeswijk Visser SJ, van der Klooster JM, Lima AP, et al. Early lactate-guided therapy in intensive care unit patients: a multicenter, open-label, randomized controlled trial. American journal of respiratory and critical care medicine. 2010;182(6):752–761.

52. Shankar-Hari M, Phillips GS, Levy ML, Seymour CW, Liu VX, Deutschman CS, et al. Developing a new definition and assessing new clinical criteria for septic shock: for the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). Jama. 2016;315(8):775–787.

53. Leone M, Asfar P, Radermacher P, Vincent JL, Martin C. Optimizing mean arterial pressure in septic shock: a critical reappraisal of the literature. Critical Care. 2015;19(1):101.

54. Harjola VP, Lassus J, Sionis A, Køber L, Tarvasmäki T, Spinar J, et al. Clinical picture and risk prediction of short-term mortality in cardiogenic shock. European journal of heart failure. 2015;17(5):501–509.

A Complementary material for the feature selection experiments

A.1 The T1 dataset

A.1.1 Feature selection results

Table 5. The *UFS* features performance measures for the *T1* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.77 ± 0.087	0.507 ± 0.179	**0.764 ± 0.195**	0.772 ± 0.125	0.768 ± 0.094
3	0.82 ± 0.075	0.58 ± 0.175	0.758 ± 0.17	0.842 ± 0.089	0.8 ± 0.091
4	0.802 ± 0.069	0.474 ± 0.203	0.54 ± 0.227	0.896 ± 0.084	0.718 ± 0.106
5	0.808 ± 0.076	0.513 ± 0.204	0.62 ± 0.224	0.876 ± 0.094	0.748 ± 0.107
6	0.798 ± 0.069	0.485 ± 0.199	0.612 ± 0.219	0.865 ± 0.077	0.738 ± 0.106
7	0.821 ± 0.082	0.541 ± 0.216	0.63 ± 0.21	0.889 ± 0.09	0.76 ± 0.11
8	0.812 ± 0.086	0.534 ± 0.203	0.658 ± 0.195	0.867 ± 0.102	0.763 ± 0.105
9	0.767 ± 0.089	0.417 ± 0.225	0.582 ± 0.212	0.834 ± 0.102	0.708 ± 0.113
10	0.798 ± 0.077	0.502 ± 0.194	0.65 ± 0.188	0.851 ± 0.087	0.75 ± 0.1
11	0.807 ± 0.074	0.536 ± 0.181	0.688 ± 0.188	0.85 ± 0.094	0.769 ± 0.094
12	0.847 ± 0.071	0.613 ± 0.188	0.676 ± 0.213	0.909 ± 0.081	0.792 ± 0.104
13	0.843 ± 0.066	0.581 ± 0.194	0.56 ± 0.228	0.944 ± 0.075	0.752 ± 0.107
14	0.838 ± 0.063	0.543 ± 0.214	0.468 ± 0.21	**0.97 ± 0.048**	0.719 ± 0.105
15	0.846 ± 0.069	0.576 ± 0.22	0.548 ± 0.239	0.952 ± 0.066	0.75 ± 0.116
16	0.837 ± 0.059	0.551 ± 0.194	0.514 ± 0.218	0.952 ± 0.061	0.733 ± 0.102
17	0.832 ± 0.067	0.526 ± 0.224	0.482 ± 0.225	0.957 ± 0.049	0.72 ± 0.113
18	0.801 ± 0.059	0.411 ± 0.228	0.356 ± 0.199	0.96 ± 0.047	0.658 ± 0.099
19	0.819 ± 0.063	0.484 ± 0.214	0.43 ± 0.195	0.959 ± 0.045	0.694 ± 0.101
20	0.821 ± 0.062	0.493 ± 0.207	0.444 ± 0.201	0.955 ± 0.053	0.699 ± 0.101
21	0.826 ± 0.068	0.517 ± 0.213	0.494 ± 0.211	0.944 ± 0.062	0.719 ± 0.106
22	0.824 ± 0.061	0.505 ± 0.2	0.464 ± 0.187	0.952 ± 0.05	0.708 ± 0.097
23	0.835 ± 0.062	0.54 ± 0.197	0.498 ± 0.203	0.955 ± 0.055	0.727 ± 0.101
24	0.856 ± 0.074	0.614 ± 0.211	0.622 ± 0.211	0.939 ± 0.063	0.781 ± 0.112
25	0.844 ± 0.073	0.584 ± 0.209	0.622 ± 0.204	0.924 ± 0.061	0.773 ± 0.109
26	0.864 ± 0.072	0.647 ± 0.198	0.692 ± 0.195	0.925 ± 0.067	0.809 ± 0.104
27	0.859 ± 0.068	0.648 ± 0.171	0.71 ± 0.18	0.913 ± 0.08	0.811 ± 0.091
28	**0.867 ± 0.07**	**0.665 ± 0.189**	0.736 ± 0.198	0.914 ± 0.071	**0.825 ± 0.102**
29	0.863 ± 0.063	0.651 ± 0.161	0.702 ± 0.164	0.921 ± 0.07	0.811 ± 0.085
30	0.847 ± 0.074	0.618 ± 0.182	0.708 ± 0.197	0.897 ± 0.088	0.803 ± 0.099

Table 6. The *RFE* features performance measures for the *T1* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.614 ± 0.208	0.0 ± 0.0	**0.26 ± 0.439**	0.74 ± 0.439	0.5 ± 0.0
3	0.737 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	**1.0 ± 0.0**	0.5 ± 0.0
4	0.734 ± 0.019	0.011 ± 0.085	0.018 ± 0.07	0.99 ± 0.035	0.504 ± 0.026
5	0.74 ± 0.04	0.079 ± 0.174	0.086 ± 0.156	0.974 ± 0.064	0.53 ± 0.064
6	0.742 ± 0.039	0.093 ± 0.174	0.108 ± 0.184	0.968 ± 0.073	0.538 ± 0.071
7	0.749 ± 0.05	0.112 ± 0.225	0.116 ± 0.212	0.975 ± 0.069	0.545 ± 0.091
8	0.742 ± 0.038	0.081 ± 0.182	0.104 ± 0.213	0.97 ± 0.069	0.537 ± 0.083
9	0.74 ± 0.069	0.093 ± 0.233	0.132 ± 0.235	0.957 ± 0.095	0.545 ± 0.104
10	0.743 ± 0.07	0.14 ± 0.25	0.194 ± 0.282	0.939 ± 0.111	0.566 ± 0.117
11	0.745 ± 0.068	0.116 ± 0.258	0.162 ± 0.28	0.953 ± 0.089	0.557 ± 0.123
12	0.743 ± 0.059	0.118 ± 0.232	0.162 ± 0.26	0.95 ± 0.085	0.556 ± 0.11
13	0.747 ± 0.058	0.15 ± 0.25	0.192 ± 0.277	0.946 ± 0.094	0.569 ± 0.113
14	0.748 ± 0.085	0.15 ± 0.275	0.19 ± 0.286	0.948 ± 0.115	0.569 ± 0.129
15	0.739 ± 0.057	0.115 ± 0.222	0.158 ± 0.258	0.947 ± 0.094	0.553 ± 0.104
16	0.753 ± 0.067	0.146 ± 0.278	0.194 ± 0.314	0.952 ± 0.086	0.573 ± 0.135
17	0.744 ± 0.06	0.12 ± 0.229	0.144 ± 0.235	0.959 ± 0.082	0.551 ± 0.103
18	0.751 ± 0.051	0.148 ± 0.241	0.186 ± 0.292	0.952 ± 0.083	0.569 ± 0.118
19	0.746 ± 0.047	0.114 ± 0.217	0.138 ± 0.241	0.963 ± 0.072	0.55 ± 0.099
20	0.739 ± 0.059	0.108 ± 0.211	0.128 ± 0.218	0.958 ± 0.089	0.543 ± 0.093
21	0.755 ± 0.074	0.182 ± 0.295	0.232 ± 0.318	0.942 ± 0.091	0.587 ± 0.14
22	0.745 ± 0.067	0.17 ± 0.261	0.218 ± 0.289	0.934 ± 0.094	0.576 ± 0.123
23	0.739 ± 0.044	0.086 ± 0.213	0.112 ± 0.218	0.964 ± 0.063	0.538 ± 0.091
24	0.751 ± 0.076	0.195 ± 0.286	0.254 ± 0.321	0.929 ± 0.106	0.591 ± 0.138
25	0.761 ± 0.072	0.221 ± 0.29	0.254 ± 0.321	0.942 ± 0.091	**0.598 ± 0.14**
26	**0.762 ± 0.062**	**0.223 ± 0.274**	0.252 ± 0.311	0.944 ± 0.088	**0.598 ± 0.13**
27	0.756 ± 0.051	0.188 ± 0.249	0.208 ± 0.28	0.951 ± 0.074	0.58 ± 0.116
28	0.751 ± 0.056	0.175 ± 0.257	0.22 ± 0.304	0.94 ± 0.092	0.58 ± 0.122
29	0.745 ± 0.052	0.156 ± 0.249	0.188 ± 0.277	0.944 ± 0.088	0.566 ± 0.111
30	0.75 ± 0.066	0.173 ± 0.257	0.19 ± 0.272	0.95 ± 0.097	0.57 ± 0.115

Table 7. The *RFECV* features performance measures for the *T1* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
1	0.642 ± 0.189	0.0 ± 0.0	0.2 ± 0.4	0.8 ± 0.4	0.5 ± 0.0
2	0.576 ± 0.224	0.0 ± 0.0	0.34 ± 0.474	0.66 ± 0.474	0.5 ± 0.0
8	0.744 ± 0.044	0.101 ± 0.199	0.124 ± 0.213	**0.966 ± 0.067**	0.545 ± 0.088
31	**0.754 ± 0.053**	**0.183 ± 0.25**	0.196 ± 0.287	0.954 ± 0.083	0.575 ± 0.117
33	0.737 ± 0.07	0.141 ± 0.258	0.202 ± 0.289	0.929 ± 0.107	0.565 ± 0.121
139	0.665 ± 0.093	0.156 ± 0.23	**0.4 ± 0.217**	0.76 ± 0.109	**0.58 ± 0.116**

The first entry uses the *recall* metric, followed by the *average precision*, the *precision*, the *weighted f1*, the *AUC* and the *neg log loss*.

Table 8. The *UFS+RFE* features performance measures for the *T1* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.753 ± 0.055	0.154 ± 0.192	0.088 ± 0.134	0.99 ± 0.099	0.539 ± 0.049
3	0.755 ± 0.034	0.165 ± 0.204	0.118 ± 0.133	0.982 ± 0.031	0.55 ± 0.063
4	0.763 ± 0.032	0.218 ± 0.199	0.142 ± 0.131	0.984 ± 0.03	0.563 ± 0.061
5	0.771 ± 0.032	0.254 ± 0.2	0.158 ± 0.134	0.99 ± 0.025	0.574 ± 0.062
6	0.769 ± 0.038	0.247 ± 0.216	0.154 ± 0.132	0.989 ± 0.026	0.571 ± 0.066
7	0.774 ± 0.032	0.287 ± 0.188	0.184 ± 0.129	0.985 ± 0.029	0.584 ± 0.06
8	0.766 ± 0.037	0.232 ± 0.212	0.15 ± 0.134	0.986 ± 0.028	0.568 ± 0.065
9	0.773 ± 0.031	0.277 ± 0.191	0.164 ± 0.118	**0.991 ± 0.024**	0.577 ± 0.057
10	0.776 ± 0.033	0.285 ± 0.199	0.174 ± 0.14	**0.991 ± 0.026**	0.582 ± 0.065
11	0.779 ± 0.045	0.314 ± 0.215	0.234 ± 0.183	0.974 ± 0.057	0.604 ± 0.081
12	0.772 ± 0.053	0.281 ± 0.232	0.222 ± 0.183	0.969 ± 0.061	0.595 ± 0.086
13	0.775 ± 0.057	0.33 ± 0.215	0.278 ± 0.213	0.953 ± 0.094	0.615 ± 0.087
14	0.774 ± 0.057	0.309 ± 0.234	0.252 ± 0.169	0.961 ± 0.059	0.606 ± 0.085
15	0.779 ± 0.054	0.33 ± 0.221	0.252 ± 0.173	0.968 ± 0.059	0.61 ± 0.083
16	0.759 ± 0.076	0.277 ± 0.235	0.226 ± 0.159	0.95 ± 0.101	0.588 ± 0.084
17	0.769 ± 0.053	0.279 ± 0.224	0.226 ± 0.176	0.963 ± 0.059	0.594 ± 0.084
18	0.772 ± 0.059	0.285 ± 0.252	0.238 ± 0.22	0.963 ± 0.07	0.6 ± 0.1
19	0.776 ± 0.053	0.317 ± 0.209	0.236 ± 0.151	0.969 ± 0.063	0.602 ± 0.074
20	0.779 ± 0.069	0.358 ± 0.207	0.28 ± 0.167	0.957 ± 0.096	0.619 ± 0.08
21	0.778 ± 0.055	0.333 ± 0.2	0.234 ± 0.15	0.973 ± 0.068	0.603 ± 0.073
22	0.779 ± 0.048	0.321 ± 0.212	0.22 ± 0.148	0.979 ± 0.044	0.6 ± 0.075
23	0.779 ± 0.035	0.303 ± 0.2	0.19 ± 0.14	0.989 ± 0.031	0.59 ± 0.065
24	0.789 ± 0.045	0.367 ± 0.191	0.246 ± 0.149	0.984 ± 0.043	0.615 ± 0.074
25	0.776 ± 0.064	0.312 ± 0.228	0.224 ± 0.161	0.973 ± 0.073	0.598 ± 0.084
26	0.777 ± 0.05	0.314 ± 0.216	0.226 ± 0.146	0.974 ± 0.051	0.6 ± 0.074
27	0.781 ± 0.056	0.353 ± 0.196	0.27 ± 0.195	0.963 ± 0.091	0.616 ± 0.08
28	0.781 ± 0.074	0.397 ± 0.231	**0.408 ± 0.273**	0.914 ± 0.129	0.661 ± 0.108
29	**0.791 ± 0.077**	**0.431 ± 0.215**	0.406 ± 0.241	0.928 ± 0.126	**0.667 ± 0.099**
30	0.778 ± 0.076	0.378 ± 0.239	0.39 ± 0.263	0.916 ± 0.113	0.653 ± 0.113

Table 9. The *RF* features performance measures for the *T1* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.831 ± 0.07	0.572 ± 0.177	0.66 ± 0.17	0.891 ± 0.082	0.776 ± 0.09
3	0.827 ± 0.072	0.562 ± 0.191	0.66 ± 0.195	0.887 ± 0.078	0.774 ± 0.1
4	**0.846 ± 0.071**	**0.615 ± 0.184**	0.713 ± 0.194	0.893 ± 0.079	0.803 ± 0.099
5	0.832 ± 0.073	0.577 ± 0.194	0.679 ± 0.205	0.886 ± 0.083	0.783 ± 0.103
6	0.814 ± 0.074	0.532 ± 0.196	0.65 ± 0.204	0.872 ± 0.084	0.761 ± 0.103
7	0.834 ± 0.074	0.608 ± 0.178	**0.763 ± 0.185**	0.86 ± 0.089	**0.811 ± 0.096**
8	0.824 ± 0.075	0.575 ± 0.186	0.72 ± 0.194	0.861 ± 0.09	0.79 ± 0.099
9	0.824 ± 0.072	0.566 ± 0.187	0.696 ± 0.206	0.869 ± 0.086	0.783 ± 0.101
10	0.812 ± 0.072	0.514 ± 0.197	0.597 ± 0.203	0.889 ± 0.081	0.743 ± 0.102
11	0.815 ± 0.075	0.521 ± 0.205	0.603 ± 0.202	0.891 ± 0.08	0.747 ± 0.104
12	0.824 ± 0.072	0.548 ± 0.193	0.635 ± 0.197	0.892 ± 0.081	0.763 ± 0.1
13	0.827 ± 0.073	0.555 ± 0.193	0.633 ± 0.196	0.896 ± 0.081	0.765 ± 0.1
14	0.829 ± 0.07	0.562 ± 0.188	0.644 ± 0.195	0.894 ± 0.082	0.769 ± 0.097
15	0.822 ± 0.071	0.54 ± 0.195	0.614 ± 0.202	**0.897 ± 0.082**	0.755 ± 0.1
16	0.818 ± 0.072	0.536 ± 0.191	0.631 ± 0.196	0.885 ± 0.083	0.758 ± 0.099
17	0.816 ± 0.075	0.535 ± 0.197	0.646 ± 0.207	0.877 ± 0.084	0.762 ± 0.104
18	0.815 ± 0.07	0.528 ± 0.19	0.634 ± 0.203	0.879 ± 0.081	0.757 ± 0.1
19	0.816 ± 0.074	0.531 ± 0.195	0.63 ± 0.202	0.882 ± 0.087	0.756 ± 0.101
20	0.806 ± 0.071	0.508 ± 0.191	0.624 ± 0.203	0.871 ± 0.085	0.747 ± 0.1
21	0.8 ± 0.076	0.49 ± 0.203	0.606 ± 0.208	0.87 ± 0.086	0.738 ± 0.105
22	0.794 ± 0.08	0.466 ± 0.217	0.569 ± 0.219	0.874 ± 0.09	0.722 ± 0.111
23	0.795 ± 0.073	0.467 ± 0.204	0.571 ± 0.216	0.875 ± 0.085	0.723 ± 0.106
24	0.796 ± 0.075	0.468 ± 0.205	0.564 ± 0.209	0.878 ± 0.084	0.721 ± 0.105
25	0.795 ± 0.076	0.468 ± 0.208	0.57 ± 0.209	0.875 ± 0.086	0.723 ± 0.106
26	0.788 ± 0.076	0.445 ± 0.217	0.553 ± 0.226	0.872 ± 0.087	0.712 ± 0.111
27	0.78 ± 0.076	0.42 ± 0.218	0.527 ± 0.217	0.87 ± 0.086	0.698 ± 0.108
28	0.781 ± 0.076	0.424 ± 0.214	0.531 ± 0.215	0.87 ± 0.087	0.7 ± 0.107
29	0.781 ± 0.075	0.423 ± 0.212	0.525 ± 0.218	0.873 ± 0.088	0.699 ± 0.107
30	0.781 ± 0.074	0.411 ± 0.223	0.501 ± 0.228	0.881 ± 0.083	0.691 ± 0.111

A.1.2 Feature sets rankings

Table 10. The *UFS* and the *RFE* feature rankings for the *T1* dataset.

R	Univariate feature selection	Recursive feature elimination
1	APACHE II_T1	E A ratio_T1=0 6
2	SOFA_T1	Aortic Valve Regurgitation_T1=Moderate
3	Respiratory rate_T1	Is the patient under other drugs_T1=Yes
4	X Norepinephrine mg kg min _T1	Sedation Scale SAS _T1
5	Number of affected organs	LA dilatation by eyeballing_T1=Yes
6	Tidal volume VT _T1	Is the patient under other drugs_T1=No
7	LA dilatation by eyeballing_T1=No	Aortic Valve Regurgitation_T1=No valv dysfunc
8	LA dilatation by eyeballing_T1=Yes	Glasgow Coma Scale_T1
9	Glasgow Coma Scale_T1	LV Hypertrophy LV mass _T1=Moderate
10	Fluid Balance ml _T1	HCO3 mmol L _T1
11	Platelet count_T1	Respiratory rate_T1
12	LV Dilatation LVEDV _T1=Moderate	LVOT average velocity time integral cm _T1
13	E A ratio_T1=1 39	SOFA_T1
14	Aortic Valve Stenosis_T1=Moderate	LA dilatation by eyeballing_T1=No
15	Base Excess mmol L _T1	LV Hypertrophy LV mass _T1=No
16	Urine Output mL day _T1	Sat O2 _T1
17	Anuria and or RRT_T1=Yes	Fibrinogen_T1 g L
18	Anuria and or RRT_T1=No	Pplat_T1
19	LVOT average diameter mm _T1	Base Excess mmol L _T1
20	Mode_T1=Volume controlled	PT_T1
21	Lactate levels mmol L _T1	Lateral e cm s _T1
22	Is the patient under other drugs_T1=Yes	Shock diagnosed within 48 from admission=No
23	Is the patient under other drugs_T1=No	Neutro abs count_T1
24	Hypotension SBP 90 mmHg or MAP 70 mmHg or SBP decrease 40 mmHg _T1=Yes	Number of affected organs
25	Site of sampling=Urine	Respiratory rate rpm _T1
26	Hypotension SBP 90 mmHg or MAP 70 mmHg or SBP decrease 40 mmHg _T1=No	Hematocrit _T1
27	Mode_T1=Pressure controlled	Heart rate bpm _T1 1
28	Sedation Scale SAS _T1	Midazolam mg kg min _T1
29	Temperature C _T1	Weight kg
30	HCO3 mmol L _T1	aPTT_T1

```
RECALL: E A ratio_T1=0 6;

AVERAGE PRECISION: Aortic Valve Regurgitation_T1=Moderate, E A ratio_T1
    =0 6;

PRECISION: Aortic Valve Regurgitation_T1=Moderate, Aortic Valve
    Regurgitation _T1=No valvular dysfunction, E A ratio_T1=0 6, Glasgow
     Coma Scale_T1, Is the patient under other drugs_T1=No, Is the
    patient under other drugs_T1=Yes, LA dilatation by eyeballing_T1=Yes
    , Sedation Scale SAS _T1;

F1 WEIGHTED: Aortic Valve Regurgitation_T1=Moderate, Aortic Valve
    Regurgitation _T1=No valvular dysfunction, Base Excess mmol L _T1, E
     A ratio_T1=0 6, Fibrinogen_T1 g L , Glasgow Coma Scale_T1, HCO3
    mmol L _T1, Heart rate bpm _T1 1, Hematocrit _T1, Is the patient
    under other drugs_T1=No, Is the patient under other drugs_T1=Yes, LA
     dilatation by eyeballing_T1=No, LA dilatation by eyeballing _T1=Yes
    , LV Hypertrophy LV mass _T1=Moderate, LV Hypertrophy LV mass _T1=No
    , LVOT average velocity time integral cm _T1, Lateral e cm s _T1,
    Midazolam mg kg min _T1, Neutro abs count_T1, Number of affected
    organs, PT_T1, Pplat_T1, Respiratory rate rpm _T1, Respiratory
    rate_T1, SOFA_T1, Sat O2 _T1, Sedation Scale SAS _T1, Value _T1, Was
     shock diagnosed within 48 from hospital admission=No, Weight kg,
    aPTT_T1;

AUC: Aortic Valve Regurgitation_T1=Moderate, Aortic Valve
    Regurgitation_T1=No valvular dysfunction, Base Excess mmol L _T1, E
    A ratio_T1=0 6, Fibrinogen_T1 g L, Glasgow Coma Scale_T1, HCO3 mmol
    L _T1, Heart rate bpm _T1 1, Hematocrit _T1, Is the patient under
    other drugs_T1=No, Is the patient under other drugs_T1=Yes, LA
    dilatation by eyeballing_T1=No, LA dilatation by eyeballing_T1=Yes,
    LV Dilatation LVEDV _T1=No, LV Hypertrophy LV mass _T1=Moderate, LV
    Hypertrophy LV mass _T1=No, LVOT average velocity time integral cm
    _T1, Lateral e cm s _T1, Midazolam mg kg min _T1, Neutro abs
    count_T1, Number of affected organs, PT_T1, Pplat_T1, Respiratory
    rate rpm _T1, Respiratory rate_T1, SOFA_T1, Sat O2 _T1, Sedation
    Scale SAS _T1, SvcO2 _T1, Value _T1, Was shock diagnosed within 48
    from hospital admission=No, Weight kg , aPTT_T1;

NEG LOG LOSS: A wave cm s _T1, AHF_T1=no, AHF_T1=yes, APACHE II_T1,
    Acute Kidney Injury_T1=AKIN I, Acute Kidney Injury_T1=AKIN II,
    Anuria and or RRT_T1=No, Anuria and or RRT_T1=Yes, Aortic Valve
    Regurgitation_T1=Mild, Aortic Valve Regurgitation_T1=Moderate,
    Aortic Valve Regurgitation_T1=No valvular assessment available,
    Aortic Valve Regurgitation_T1=No valvular dysfunction, Aortic Valve
    Stenosis_T1=Moderate, Aortic Valve Stenosis_T1=No valvular
    assessment available, BMI, Base Excess mmol L _T1, Bilirubin mg dL
    _T1, Chloride_T1, Creatinine mg dL _T1, Diastolic Blood Pressure
    mmHg _T1, E A ratio_T1=0.6, E A ratio_T1=0.62, E A ratio_T1=0.69, E
    A ratio_T1=0.72, E A ratio_T1=0.73, E A ratio_T1=0.78, E A ratio_T1
    =0.81, E A ratio_T1=0.84, E A ratio_T1=0.89, E A ratio_T1=0.97, E A
    ratio_T1=1, E A ratio_T1=1.01, E A ratio_T1=1.18, E A ratio_T1=1.3,
```

```
    E A ratio_T1=9, E e _T1, E wave cm s _T1, Fibrinogen_T1 g L , Gender
    =Female, Gender=Male, Glasgow Coma Scale_T1, Glycemia mg dL _T1,
    HCO3 mmol L _T1, Heart rate bpm _T1, Heart rate bpm _T1 1,
    Hematocrit _T1, Hypotension SBP 90 mmHg or MAP 70 mmHg or SBP
    decrease 40 mmHg _T1=No, Hypotension SBP 90 mmHg or MAP 70 mmHg or
    SBP decrease 40 mmHg _T1=Yes, Is the patient affected by Acute
    Myocardial Infarction AMI _T1=No, Is the patient affected by Acute
    Myocardial Infarction AMI _T1=Yes, Is the patient affected by
    Prolonged Arrhythmias PA _T1=No, Is the patient affected by
    Prolonged Arrhythmias PA _T1=Yes, Is the patient under inotropic
    drugs_T1=No, Is the patient under inotropic drugs_T1=Yes, Is the
    patient under other drugs_T1=No, Is the patient under other drugs_T1
    =Yes, Is the patient under sedation drugs_T1=No, Is the patient
    under sedation drugs_T1=Yes, LA dilatation by eyeballing_T1=No, LA
    dilatation by eyeballing_T1=Yes, LV Dilatation LVEDV _T1=No, LV
    Dilatation LVEDV _T1=Severe, LV Hypertrophy LV mass _T1=Mild, LV
    Hypertrophy LV mass _T1=Moderate, LV Hypertrophy LV mass _T1=No,
    LVOT average diameter mm _T1, LVOT average velocity time integral cm
     _T1, Lactate levels mmol L _T1, Lateral e cm s _T1, Leukocytes
    total 1000 mm 3 _T1, Lympho abs count_T1, Mean arterial pressure
    mmHg _T1, Midazolam \xb5g kg min _T1, Mitral Valve Regurgitation_T1=
    Mild, Mitral Valve Regurgitation_T1=No valvular assessment available
    , Mitral Valve Regurgitation_T1=No valvular dysfunction, ...
... Mode_T1=Full pressure support, Mode_T1=Pressure controlled, Mode_T1=
    Volume controlled, Na mmol L _T1, Neutro abs count_T1, Number of
    affected organs, PT_T1, PaO2 FiO2_T1, Paradoxical interventricular
    septum movements_T1=Moderate, Paradoxical interventricular septum
    movements_T1=No, Paradoxical interventricular septum movements_T1=
    Severe, Platelet count_T1, Platelets 10 3 mm 3 _T1, Pplat_T1, RBC
    count_T1, RV cross section D shape_T1=No, RV cross section D
    shape_T1=Yes, Renal Replacement Therapy_T1=No, Renal Replacement
    Therapy_T1=Yes, Respiratory rate rpm _T1, Respiratory rate_T1,
    Result=Negative, Result=Positive, SOFA_T1, Sample 1=Negative, Sample
     1=Positive, Sample 2=Negative, Sample 2=Positive, Sample 3=Negative
    , Sample 3=Positive, Sat O2 _T1, Sedation Scale SAS _T1, Segmental
    LV wall kinetics_T1=Normal, Segmental LV wall kinetics_T1=Reduced,
    Site of sampling 1=Abdominal drain, Site of sampling 1=Other, Site
    of sampling 1=Respiratory tract, Site of sampling 2=Respiratory
    tract, Site of sampling 2=Urine, Site of sampling=Abdominal drain,
    Site of sampling=CSF, Site of sampling=Respiratory tract, Site of
    sampling=Urine, SvcO2 _T1, Systolic blood pressure mmHg _T1,
    Temperature C _T1, Tidal volume VT _T1, Tracheal Intubation_T1=No,
    Tracheal Intubation_T1=Yes, Tricuspid Valve Regurgitation_T1=Mild,
    Tricuspid Valve Regurgitation_T1=No valvular assessment available,
    Tricuspid Valve Regurgitation_T1=No valvular dysfunction, Tricuspid
    annular plane systolic excursion TAPSE mm _T1, Value _T1, WBC abs
    count_T1, Was shock diagnosed within 48 from hospital admission=No,
    Was shock diagnosed within 48 from hospital admission=Yes, Was the
    patient transfused_T1=No, Was the patient transfused_T1=Yes, Weight
    kg, Which type of shock=Cardiogenic, Which type of shock=Septic,
    aPTT_T1;
```

List of features for the *RFECV* with the *average precision* and the *accuracy* metrics for

the **T1** dataset.

Table 11. The *UFS+RFE* and the *RF* feature rankings for the *T1* dataset.

R	UFS+FSE	Random Forest
1	E A ratio_T1=1 39	APACHE II_T1
2	Site of sampling=CSF	Urine Output mL day _T1
3	E A ratio_T1=0 69	Respiratory rate_T1
4	Aortic Valve Stenosis_T1=Moderate	E wave cm s _T1
5	Mode_T1=Volume controlled	Lactate levels mmol L _T1
6	LA dilatation by eyeballing_T1=Yes	Tidal volume VT _T1
7	RBC count_T1	SOFA_T1
8	Is the patient under other drugs_-T1=Yes	Fluid Balance ml _T1
9	Is the patient under sedation drugs_-T1=No	E e _T1
10	Is the patient under other drugs_-T1=No	X Norepinephrine mg kg min _T1
11	E A ratio_T1=0 81	Heart rate bpm _T1
12	Is the patient under sedation drugs_-T1=Yes	Pplat_T1
13	Glasgow Coma Scale_T1	Temperature C _T1
14	LV Hypertrophy LV mass _-T1=Moderate	pH_T1
15	SOFA_T1	Fibrinogen_T1 g L
16	Sedation Scale SAS _T1	Respiratory rate rpm _T1
17	Aortic Valve Regurgitation_T1=No val dysfunc	Platelet count_T1
18	HCO3 mmol L _T1	PEEP_T1
19	pH_T1	Leukocytes total 1000 mm 3 _T1
20	FiO2_T1	LVOT av vel time integral cm _T1
21	LV Dilatation LVEDV _T1=Moderate	PaO2 FiO2_T1
22	Aortic Valve Regurgitation_T1=Severe	Midazolam mg kg min _T1
23	Respiratory rate_T1	Base Excess mmol L _T1
24	Is the patient under inotropic drugs_-T1=No	Neutro abs count_T1
25	LA dilatation by eyeballing_T1=No	A wave cm s _T1
26	LV Dilatation LVEDV _T1=No	Glycemia mg dL _T1
27	Hypotension SBP 90 mmHg or MAP 70 mmHg or SBP decrease 40 mmHg _T1=Yes	Platelets 10 3 mm 3 _T1
28	Hypotension SBP 90 mmHg or MAP 70 mmHg or SBP decrease 40 mmHg _T1=No	RBC count_T1
29	Mean arterial pressure mmHg _T1	LVOT average diameter mm _T1
30	Is the patient under inotropic drugs_-T1=Yes	Lympho abs count_T1

A.2 The T1+T2 dataset

A.2.1 Feature selection results

Table 12. The *UFS* features performance measures for the *T1+T2* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.829 ± 0.074	0.567 ± 0.197	0.662 ± 0.201	0.889 ± 0.078	0.776 ± 0.104
3	0.805 ± 0.072	0.549 ± 0.157	0.738 ± 0.162	0.829 ± 0.097	0.784 ± 0.081
4	0.776 ± 0.088	0.489 ± 0.202	0.72 ± 0.188	0.796 ± 0.099	0.758 ± 0.106
5	0.785 ± 0.088	0.477 ± 0.231	0.66 ± 0.231	0.829 ± 0.1	0.745 ± 0.119
6	0.789 ± 0.069	0.465 ± 0.183	0.598 ± 0.182	0.858 ± 0.076	0.728 ± 0.095
7	0.823 ± 0.077	0.548 ± 0.211	0.646 ± 0.226	0.886 ± 0.083	0.766 ± 0.113
8	0.826 ± 0.071	0.57 ± 0.177	0.672 ± 0.213	0.881 ± 0.098	0.777 ± 0.098
9	0.802 ± 0.094	0.537 ± 0.223	0.726 ± 0.213	0.829 ± 0.103	0.777 ± 0.118
10	0.784 ± 0.098	0.525 ± 0.19	0.76 ± 0.167	0.793 ± 0.124	0.776 ± 0.1
11	0.804 ± 0.104	0.563 ± 0.206	0.768 ± 0.178	0.816 ± 0.128	0.792 ± 0.108
12	0.831 ± 0.091	0.602 ± 0.189	0.748 ± 0.159	0.861 ± 0.109	0.804 ± 0.097
13	0.831 ± 0.091	0.605 ± 0.187	0.76 ± 0.162	0.856 ± 0.111	0.808 ± 0.097
14	0.839 ± 0.086	0.628 ± 0.196	**0.798 ± 0.175**	0.854 ± 0.091	**0.826 ± 0.104**
15	0.827 ± 0.088	0.606 ± 0.186	0.79 ± 0.163	0.84 ± 0.104	0.815 ± 0.097
16	0.818 ± 0.082	0.581 ± 0.182	0.766 ± 0.172	0.836 ± 0.098	0.801 ± 0.095
17	0.807 ± 0.1	0.557 ± 0.213	0.738 ± 0.18	0.832 ± 0.115	0.785 ± 0.11
18	0.818 ± 0.08	0.568 ± 0.186	0.72 ± 0.177	0.854 ± 0.096	0.787 ± 0.096
19	0.813 ± 0.085	0.568 ± 0.204	0.762 ± 0.197	0.831 ± 0.089	0.797 ± 0.11
20	0.813 ± 0.083	0.559 ± 0.191	0.73 ± 0.171	0.842 ± 0.096	0.786 ± 0.097
21	0.812 ± 0.076	0.558 ± 0.171	0.742 ± 0.17	0.836 ± 0.089	0.789 ± 0.092
22	0.81 ± 0.076	0.551 ± 0.187	0.734 ± 0.194	0.837 ± 0.083	0.786 ± 0.102
23	0.789 ± 0.089	0.5 ± 0.21	0.682 ± 0.192	0.828 ± 0.097	0.755 ± 0.109
24	0.838 ± 0.094	0.626 ± 0.215	0.796 ± 0.179	0.854 ± 0.097	0.825 ± 0.111
25	0.846 ± 0.067	0.631 ± 0.166	0.774 ± 0.164	0.871 ± 0.076	0.823 ± 0.086
26	**0.867 ± 0.076**	**0.667 ± 0.193**	0.738 ± 0.187	0.914 ± 0.077	**0.826 ± 0.102**
27	0.866 ± 0.075	0.664 ± 0.186	0.734 ± 0.179	0.913 ± 0.078	0.823 ± 0.098
28	0.865 ± 0.067	0.659 ± 0.177	0.722 ± 0.192	**0.916 ± 0.073**	0.819 ± 0.097
29	0.852 ± 0.076	0.63 ± 0.202	0.724 ± 0.193	0.898 ± 0.078	0.811 ± 0.103
30	0.849 ± 0.068	0.623 ± 0.175	0.71 ± 0.186	0.899 ± 0.076	0.805 ± 0.094

Table 13. The *RFE* features performance measures for the *T1+T2* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.838 ± 0.071	0.585 ± 0.182	0.638 ± 0.197	**0.909 ± 0.086**	0.774 ± 0.097
3	**0.841 ± 0.068**	**0.603 ± 0.179**	0.698 ± 0.195	0.891 ± 0.082	0.795 ± 0.096
4	0.835 ± 0.08	0.589 ± 0.209	0.694 ± 0.201	0.886 ± 0.087	0.79 ± 0.106
5	0.825 ± 0.077	0.576 ± 0.18	0.706 ± 0.178	0.867 ± 0.094	0.787 ± 0.094
6	0.802 ± 0.067	0.499 ± 0.167	0.622 ± 0.165	0.866 ± 0.079	0.744 ± 0.086
7	0.818 ± 0.078	0.537 ± 0.2	0.634 ± 0.188	0.884 ± 0.087	0.759 ± 0.101
8	0.828 ± 0.077	0.596 ± 0.189	0.766 ± 0.204	0.85 ± 0.092	0.808 ± 0.102
9	0.799 ± 0.079	0.575 ± 0.157	**0.834 ± 0.15**	0.786 ± 0.101	0.81 ± 0.084
10	0.797 ± 0.084	0.56 ± 0.183	0.8 ± 0.17	0.796 ± 0.102	0.798 ± 0.095
11	0.812 ± 0.079	0.586 ± 0.179	0.816 ± 0.169	0.811 ± 0.087	**0.813 ± 0.096**
12	0.786 ± 0.094	0.523 ± 0.205	0.754 ± 0.192	0.798 ± 0.11	0.776 ± 0.109
13	0.785 ± 0.082	0.499 ± 0.193	0.702 ± 0.197	0.815 ± 0.097	0.758 ± 0.103
14	0.805 ± 0.073	0.538 ± 0.183	0.718 ± 0.204	0.836 ± 0.091	0.777 ± 0.099
15	0.791 ± 0.071	0.503 ± 0.169	0.69 ± 0.191	0.826 ± 0.09	0.758 ± 0.093
16	0.786 ± 0.079	0.498 ± 0.19	0.692 ± 0.203	0.82 ± 0.1	0.756 ± 0.101
17	0.805 ± 0.076	0.534 ± 0.175	0.704 ± 0.182	0.841 ± 0.092	0.773 ± 0.095
18	0.79 ± 0.074	0.505 ± 0.198	0.706 ± 0.216	0.82 ± 0.086	0.763 ± 0.106
19	0.793 ± 0.078	0.509 ± 0.189	0.688 ± 0.195	0.831 ± 0.095	0.759 ± 0.1
20	0.792 ± 0.075	0.52 ± 0.178	0.728 ± 0.197	0.815 ± 0.096	0.771 ± 0.097
21	0.779 ± 0.082	0.494 ± 0.199	0.722 ± 0.215	0.799 ± 0.101	0.761 ± 0.107
22	0.797 ± 0.083	0.524 ± 0.194	0.718 ± 0.2	0.825 ± 0.097	0.772 ± 0.105
23	0.792 ± 0.083	0.524 ± 0.2	0.744 ± 0.208	0.809 ± 0.094	0.776 ± 0.109
24	0.778 ± 0.093	0.488 ± 0.212	0.702 ± 0.201	0.805 ± 0.108	0.754 ± 0.111
25	0.785 ± 0.087	0.487 ± 0.227	0.682 ± 0.232	0.822 ± 0.096	0.752 ± 0.119
26	0.807 ± 0.084	0.542 ± 0.21	0.71 ± 0.225	0.841 ± 0.099	0.776 ± 0.113
27	0.783 ± 0.077	0.477 ± 0.196	0.654 ± 0.217	0.829 ± 0.095	0.742 ± 0.105
28	0.778 ± 0.092	0.463 ± 0.226	0.646 ± 0.226	0.825 ± 0.106	0.735 ± 0.119
29	0.784 ± 0.092	0.472 ± 0.228	0.638 ± 0.222	0.836 ± 0.1	0.737 ± 0.12
30	0.773 ± 0.075	0.453 ± 0.197	0.64 ± 0.224	0.82 ± 0.101	0.73 ± 0.104

Table 14. The *RFECV* features performance measures for the *T1+T2* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
3	**0.838 ± 0.07**	**0.588 ± 0.183**	**0.668 ± 0.19**	**0.899 ± 0.078**	**0.783 ± 0.097**
30	0.782 ± 0.078	0.473 ± 0.21	0.65 ± 0.257	0.829 ± 0.105	0.74 ± 0.116

The first feature set was achieved by using the *accuracy* and the second one with the *average precision* metric.

Table 15. The *UFS+RFE* features performance measures for the *T1+T2* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.771 ± 0.059	0.275 ± 0.198	0.156 ± 0.14	**0.99 ± 0.099**	0.573 ± 0.056
3	0.768 ± 0.034	0.26 ± 0.193	0.168 ± 0.119	0.982 ± 0.031	0.575 ± 0.058
4	0.763 ± 0.034	0.227 ± 0.2	0.156 ± 0.131	0.979 ± 0.032	0.568 ± 0.061
5	0.779 ± 0.031	0.324 ± 0.166	0.214 ± 0.124	0.981 ± 0.032	0.597 ± 0.057
6	0.776 ± 0.041	0.301 ± 0.207	0.204 ± 0.147	0.981 ± 0.032	0.592 ± 0.072
7	0.784 ± 0.044	0.326 ± 0.215	0.21 ± 0.156	0.989 ± 0.026	0.599 ± 0.078
8	0.783 ± 0.038	0.329 ± 0.196	0.212 ± 0.141	0.986 ± 0.028	0.599 ± 0.069
9	0.781 ± 0.04	0.312 ± 0.206	0.2 ± 0.144	0.988 ± 0.027	0.594 ± 0.071
10	0.783 ± 0.037	0.339 ± 0.184	0.21 ± 0.124	0.988 ± 0.027	0.599 ± 0.063
11	**0.785 ± 0.041**	0.34 ± 0.199	0.228 ± 0.152	0.984 ± 0.03	0.606 ± 0.074
12	0.775 ± 0.036	0.295 ± 0.193	0.2 ± 0.133	0.98 ± 0.032	0.59 ± 0.063
13	0.776 ± 0.038	0.31 ± 0.186	0.21 ± 0.134	0.979 ± 0.033	0.594 ± 0.065
14	0.779 ± 0.037	0.313 ± 0.195	0.21 ± 0.142	0.983 ± 0.031	0.596 ± 0.068
15	0.776 ± 0.034	0.303 ± 0.185	0.196 ± 0.13	0.984 ± 0.03	0.59 ± 0.062
16	0.78 ± 0.04	0.317 ± 0.2	0.212 ± 0.141	0.983 ± 0.031	0.597 ± 0.069
17	0.773 ± 0.04	0.285 ± 0.206	0.202 ± 0.146	0.977 ± 0.033	0.59 ± 0.07
18	0.776 ± 0.043	0.32 ± 0.186	0.24 ± 0.139	0.967 ± 0.046	0.604 ± 0.067
19	0.774 ± 0.04	0.296 ± 0.2	0.228 ± 0.155	0.969 ± 0.038	0.599 ± 0.072
20	0.765 ± 0.041	0.266 ± 0.203	0.21 ± 0.158	0.964 ± 0.046	0.587 ± 0.073
21	0.775 ± 0.043	0.31 ± 0.2	0.232 ± 0.149	0.969 ± 0.042	0.6 ± 0.072
22	0.774 ± 0.038	0.308 ± 0.178	0.244 ± 0.143	0.963 ± 0.042	0.603 ± 0.066
23	0.765 ± 0.043	0.253 ± 0.214	0.198 ± 0.154	0.968 ± 0.042	0.583 ± 0.073
24	0.77 ± 0.04	0.29 ± 0.188	0.218 ± 0.136	0.967 ± 0.043	0.593 ± 0.065
25	0.778 ± 0.039	0.315 ± 0.195	0.226 ± 0.151	0.976 ± 0.035	0.601 ± 0.071
26	0.769 ± 0.041	0.27 ± 0.207	0.216 ± 0.167	0.967 ± 0.045	0.592 ± 0.076
27	0.774 ± 0.043	0.297 ± 0.208	0.24 ± 0.165	0.964 ± 0.04	0.602 ± 0.078
28	**0.785 ± 0.043**	**0.342 ± 0.211**	**0.262 ± 0.178**	0.972 ± 0.043	**0.617 ± 0.081**
29	0.776 ± 0.044	0.301 ± 0.211	0.248 ± 0.17	0.965 ± 0.038	0.607 ± 0.081
30	0.781 ± 0.047	0.328 ± 0.209	0.256 ± 0.165	0.968 ± 0.041	0.612 ± 0.08

Table 16. The *RF* performance measures for the *T1+T2* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.832 ± 0.064	0.552 ± 0.185	0.572 ± 0.21	0.925 ± 0.072	0.749 ± 0.1
3	0.833 ± 0.08	0.569 ± 0.211	0.64 ± 0.219	0.902 ± 0.081	0.771 ± 0.114
4	0.855 ± 0.075	0.638 ± 0.187	0.73 ± 0.166	0.9 ± 0.076	0.815 ± 0.095
5	0.854 ± 0.075	0.646 ± 0.183	**0.768 ± 0.169**	0.884 ± 0.083	**0.826 ± 0.094**
6	0.848 ± 0.067	0.629 ± 0.172	0.746 ± 0.194	0.885 ± 0.074	0.816 ± 0.097
7	0.856 ± 0.057	0.634 ± 0.148	0.68 ± 0.167	0.919 ± 0.073	0.8 ± 0.08
8	0.851 ± 0.066	0.624 ± 0.171	0.702 ± 0.18	0.904 ± 0.074	0.803 ± 0.092
9	**0.864 ± 0.064**	**0.658 ± 0.166**	0.718 ± 0.188	0.916 ± 0.075	0.817 ± 0.092
10	0.847 ± 0.072	0.605 ± 0.191	0.658 ± 0.199	0.914 ± 0.074	0.786 ± 0.103
11	0.847 ± 0.067	0.618 ± 0.177	0.694 ± 0.213	0.902 ± 0.086	0.798 ± 0.099
12	0.84 ± 0.07	0.605 ± 0.173	0.696 ± 0.184	0.891 ± 0.088	0.794 ± 0.091
13	0.852 ± 0.07	0.637 ± 0.179	0.74 ± 0.199	0.891 ± 0.084	0.816 ± 0.097
14	0.844 ± 0.075	0.618 ± 0.19	0.728 ± 0.195	0.886 ± 0.087	0.807 ± 0.1
15	0.844 ± 0.08	0.622 ± 0.191	0.736 ± 0.179	0.883 ± 0.091	0.809 ± 0.098
16	0.831 ± 0.072	0.581 ± 0.171	0.686 ± 0.186	0.883 ± 0.09	0.784 ± 0.093
17	0.828 ± 0.078	0.582 ± 0.186	0.702 ± 0.191	0.874 ± 0.094	0.788 ± 0.098
18	0.833 ± 0.073	0.588 ± 0.187	0.698 ± 0.207	0.881 ± 0.09	0.789 ± 0.101
19	0.84 ± 0.07	0.604 ± 0.175	0.692 ± 0.199	0.893 ± 0.087	0.792 ± 0.097
20	0.832 ± 0.072	0.593 ± 0.18	0.73 ± 0.199	0.869 ± 0.081	0.799 ± 0.1
21	0.845 ± 0.069	0.629 ± 0.167	0.764 ± 0.184	0.874 ± 0.082	0.819 ± 0.093
22	0.837 ± 0.077	0.601 ± 0.194	0.724 ± 0.191	0.878 ± 0.083	0.801 ± 0.102
23	0.832 ± 0.081	0.589 ± 0.2	0.698 ± 0.203	0.88 ± 0.096	0.789 ± 0.105
24	0.839 ± 0.072	0.605 ± 0.192	0.732 ± 0.202	0.877 ± 0.083	0.805 ± 0.101
25	0.828 ± 0.073	0.562 ± 0.174	0.608 ± 0.194	0.906 ± 0.097	0.757 ± 0.093
26	0.828 ± 0.071	0.556 ± 0.184	0.594 ± 0.205	0.911 ± 0.092	0.753 ± 0.097
27	0.831 ± 0.06	0.538 ± 0.194	0.562 ± 0.219	**0.926 ± 0.067**	0.744 ± 0.102
28	0.826 ± 0.074	0.547 ± 0.204	0.582 ± 0.212	0.913 ± 0.093	0.747 ± 0.102
29	0.822 ± 0.073	0.534 ± 0.199	0.592 ± 0.212	0.904 ± 0.083	0.748 ± 0.105
30	0.821 ± 0.067	0.524 ± 0.205	0.58 ± 0.222	0.907 ± 0.07	0.744 ± 0.107

A.2.2 Feature sets rankings

Table 17. The *UFS* and *RFE* feature rankings for the *T1+T2* dataset.

R	Univariate feature selection	Recursive feature elimination
1	APACHE II_T1	A wave cm s _T1
2	Anuria and or RRT_T1=Yes	APACHE II_T1
3	Aortic Valve Stenosis_T1=Moderate	Base Excess mmol L _T1
4	Base Excess mmol L _T1	Glasgow Coma Scale_T1
5	Creatinine mg dL _T2	Glasgow Coma Scale_T2
6	E A ratio_T1=1.39	HCO3 mmol L _T1
7	FiO2_T2	Heart rate bpm _T1
8	Fluid Balance ml _T1	Heart rate bpm _T1 1
9	Glasgow Coma Scale_T1	Heart rate bpm _T2 1
10	Glasgow Coma Scale_T2	Hematocrit _T1
11	K mmol L _T2	Hematocrit _T2
12	LA dilatation by eyeballing_T1=No	LVOT average diameter mm _T1
13	LA dilatation by eyeballing_T1=Yes	Leukocytes total 1000 mm 3 _T2
14	LV Dilatation LVEDV _T1=Moderate	Na mmol L _T1
15	Lactate levels mmol L _T2	PEEP_T2
16	Number of affected organs	PT_T1
17	Platelet count_T1	PT_T2
18	Platelet count_T2	PaO2 mmHg _T1
19	RBC count_T2	PaO2 mmHg _T2
20	Respiratory rate_T1	Platelet count_T1
21	SOFA_T1	Pplat_T2
22	SOFA_T2	Respiratory rate rpm _T1
23	Sat O2 FiO2_T2	Respiratory rate_T2
24	Sedation Scale SAS _T2	SOFA_T1
25	Tidal volume VT _T1	SOFA_T2
26	Tracheal Intubation_T2=No	SvcO2 _T1
27	Tracheal Intubation_T2=Yes	Tidal volume VT _T1
28	Urine Output mL day _T1	WBC abs count_T2
29	X Norepinephrine μg kg min _T1	Weight kg
30	pH_T2	aPTT_T2

```
ACCURACY: Glasgow Coma Scale_T1, LVOT average diameter mm _T1, SOFA_T2;

AVERAGE PRECISION: A wave cm s _T1, APACHE II_T1, Base Excess mmol L _T1
    , Glasgow Coma Scale_T1, Glasgow Coma Scale_T2, HCO3 mmol L _T1,
    Heart rate bpm _T1, Heart rate bpm _T1 1, Heart rate bpm _T2 1,
    Hematocrit _T1, Hematocrit _T2, LVOT average diameter mm _T1,
    Leukocytes total 1000 mm 3 _T2, Na mmol L _T1, PEEP_T2, PT_T1, PT_T2
    , PaO2 mmHg _T1, PaO2 mmHg _T2, Platelet count_T1, Pplat_T2,
    Respiratory rate rpm _T1, Respiratory rate_T2, SOFA_T1, SOFA_T2,
    SvcO2 _T1, Tidal volume VT _T1, WBC abs count_T2, Weight kg ,
    aPTT_T2.
```

List of features for the *RFECV* with the *accuracy* and the *average precision* metrics for the **T1+T2** dataset.

Table 18. The *UFS+RFE* and *RF* feature rankings for the *T1+T2* dataset.

R	UFS+RFE	Random Forest
1	E A ratio_T1=1.39	SOFA_T2
2	LV Dilatation LVEDV _T1=Moderate	Lactate levels mmol L _T2
3	E A ratio_T1=0.69	APACHE II_T1
4	E A ratio_T1=0.6	Respiratory rate_T1
5	Aortic Valve Stenosis_T1=Moderate	Urine Output mL day _T1
6	Is the patient under other drugs_-T1=Yes	Urine Output mL day _T2
7	Sedation Scale SAS _T2	X Norepinephrine μg kg min _T1
8	Site of sampling=CSF	E wave cm s _T1
9	Number of affected organs	pH_T2
10	Is the patient under other drugs_-T1=No	Pplat_T2
11	Site of sampling=Urine	SOFA_T1
12	Aortic Valve Regurgitation_-T2=Moderate	Tidal volume VT _T1
13	Aortic Valve Regurgitation_-T1=Moderate	Fluid Balance ml _T2
14	Is the patient under sedation drugs_-T2=No	Glasgow Coma Scale_T2
15	K mmol L _T2	Creatinine mg dL _T2
16	HCO3 mmol L _T1	Lactate levels mmol L _T1
17	Segments affected by eyeballing_T2	Fluid Balance ml _T1
18	E A ratio_T1=0.76	Respiratory rate_T2
19	Glasgow Coma Scale_T1	Heart rate bpm _T1
20	Temperature C _T1	E e _T1
21	Aortic Valve Regurgitation_T1=Severe	PT_T2
22	Is the patient under sedation drugs_-T2=Yes	FiO2_T2
23	Acute Kidney Injury_T2=AKIN III	Platelet count_T2
24	SOFA_T1	Sat O2 _T2
25	Tricuspid Valve Regurgitation_-T2=Mild	Lympho abs count_T2
26	Base Excess mmol L _T1	PaO2 mmHg _T2
27	Anuria and or RRT_T1=No	PEEP_T2
28	Anuria and or RRT_T1=Yes	Temperature C _T1
29	Respiratory rate_T1	pH_T1
30	pH_T2	Pplat_T1

A.3 The Full dataset

A.3.1 Feature selection results

Table 19. The *UFS* features performance measures for the *Full* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.831 ± 0.073	0.558 ± 0.207	0.634 ± 0.198	**0.901 ± 0.072**	0.767 ± 0.104
3	0.847 ± 0.059	0.616 ± 0.16	0.71 ± 0.191	0.896 ± 0.07	0.803 ± 0.09
4	0.856 ± 0.061	0.64 ± 0.162	0.734 ± 0.204	0.899 ± 0.07	0.817 ± 0.097
5	0.863 ± 0.069	0.662 ± 0.175	0.754 ± 0.187	**0.901 ± 0.078**	0.828 ± 0.095
6	0.853 ± 0.072	0.656 ± 0.16	0.806 ± 0.14	0.87 ± 0.084	0.838 ± 0.081
7	0.861 ± 0.083	**0.681 ± 0.184**	**0.834 ± 0.15**	0.87 ± 0.094	**0.852 ± 0.092**
8	0.853 ± 0.069	0.65 ± 0.165	0.782 ± 0.181	0.879 ± 0.081	0.83 ± 0.092
9	0.864 ± 0.079	0.674 ± 0.194	0.806 ± 0.18	0.884 ± 0.079	0.845 ± 0.102
10	0.848 ± 0.078	0.638 ± 0.178	0.766 ± 0.172	0.877 ± 0.094	0.822 ± 0.093
11	0.845 ± 0.078	0.624 ± 0.187	0.742 ± 0.192	0.882 ± 0.086	0.812 ± 0.102
12	0.838 ± 0.071	0.585 ± 0.188	0.674 ± 0.193	0.897 ± 0.072	0.786 ± 0.101
13	0.834 ± 0.069	0.573 ± 0.201	0.67 ± 0.214	0.892 ± 0.071	0.781 ± 0.106
14	0.831 ± 0.07	0.583 ± 0.177	0.696 ± 0.184	0.879 ± 0.084	0.788 ± 0.093
15	0.832 ± 0.074	0.578 ± 0.197	0.696 ± 0.207	0.88 ± 0.076	0.788 ± 0.107
16	0.835 ± 0.06	0.587 ± 0.152	0.684 ± 0.179	0.889 ± 0.08	0.786 ± 0.084
17	0.836 ± 0.07	0.593 ± 0.186	0.718 ± 0.19	0.878 ± 0.071	0.798 ± 0.1
18	0.84 ± 0.071	0.602 ± 0.19	0.72 ± 0.196	0.883 ± 0.074	0.801 ± 0.101
19	**0.869 ± 0.068**	**0.681 ± 0.171**	0.796 ± 0.167	0.895 ± 0.072	0.845 ± 0.09
20	0.861 ± 0.08	0.663 ± 0.194	0.772 ± 0.183	0.892 ± 0.087	0.832 ± 0.101
21	0.839 ± 0.065	0.609 ± 0.16	0.738 ± 0.167	0.875 ± 0.078	0.806 ± 0.085
22	0.834 ± 0.093	0.597 ± 0.222	0.728 ± 0.182	0.871 ± 0.103	0.8 ± 0.109
23	0.841 ± 0.076	0.615 ± 0.176	0.744 ± 0.155	0.876 ± 0.085	0.81 ± 0.089
24	0.848 ± 0.066	0.619 ± 0.169	0.706 ± 0.175	0.899 ± 0.074	0.803 ± 0.09
25	0.84 ± 0.078	0.615 ± 0.181	0.734 ± 0.17	0.878 ± 0.092	0.806 ± 0.093
26	0.838 ± 0.085	0.601 ± 0.204	0.714 ± 0.184	0.882 ± 0.091	0.798 ± 0.105
27	0.821 ± 0.073	0.562 ± 0.181	0.684 ± 0.192	0.87 ± 0.092	0.777 ± 0.095
28	0.816 ± 0.076	0.554 ± 0.184	0.694 ± 0.182	0.86 ± 0.091	0.777 ± 0.096
29	0.823 ± 0.085	0.566 ± 0.213	0.692 ± 0.199	0.87 ± 0.094	0.781 ± 0.109
30	0.823 ± 0.08	0.567 ± 0.195	0.708 ± 0.184	0.864 ± 0.082	0.786 ± 0.104

Table 20. The *RFE* features performance measures for the *Full* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.719 ± 0.036	0.006 ± 0.156	0.068 ± 0.155	**0.951 ± 0.059**	0.51 ± 0.063
3	0.738 ± 0.053	0.137 ± 0.23	0.142 ± 0.161	**0.951 ± 0.061**	0.546 ± 0.078
4	0.738 ± 0.059	0.17 ± 0.234	0.208 ± 0.237	0.928 ± 0.086	0.568 ± 0.101
5	0.755 ± 0.054	0.258 ± 0.215	0.268 ± 0.186	0.929 ± 0.072	0.598 ± 0.083
6	0.752 ± 0.061	0.248 ± 0.237	0.28 ± 0.212	0.92 ± 0.076	0.6 ± 0.096
7	0.744 ± 0.064	0.222 ± 0.24	0.24 ± 0.174	0.924 ± 0.071	0.582 ± 0.089
8	0.747 ± 0.065	0.266 ± 0.229	0.314 ± 0.228	0.901 ± 0.1	0.608 ± 0.097
9	0.764 ± 0.068	0.307 ± 0.26	0.346 ± 0.24	0.913 ± 0.078	0.629 ± 0.112
10	0.725 ± 0.073	0.194 ± 0.228	0.262 ± 0.193	0.891 ± 0.102	0.576 ± 0.091
11	0.751 ± 0.073	0.267 ± 0.248	0.33 ± 0.234	0.901 ± 0.081	0.615 ± 0.112
12	0.729 ± 0.068	0.195 ± 0.24	0.268 ± 0.219	0.894 ± 0.096	0.581 ± 0.098
13	0.723 ± 0.088	0.212 ± 0.251	0.312 ± 0.22	0.869 ± 0.111	0.591 ± 0.111
14	0.737 ± 0.077	0.247 ± 0.236	0.326 ± 0.203	0.884 ± 0.09	0.605 ± 0.103
15	0.731 ± 0.071	0.254 ± 0.216	0.354 ± 0.215	0.866 ± 0.103	0.61 ± 0.096
16	0.749 ± 0.078	0.286 ± 0.248	0.352 ± 0.206	0.891 ± 0.089	0.621 ± 0.106
17	0.747 ± 0.081	0.279 ± 0.265	0.36 ± 0.223	0.886 ± 0.085	0.623 ± 0.115
18	0.741 ± 0.082	0.285 ± 0.238	0.37 ± 0.201	0.874 ± 0.097	0.622 ± 0.104
19	0.757 ± 0.08	0.339 ± 0.232	0.434 ± 0.206	0.873 ± 0.082	0.653 ± 0.11
20	0.742 ± 0.087	0.291 ± 0.234	0.376 ± 0.173	0.872 ± 0.096	0.624 ± 0.101
21	0.747 ± 0.076	0.3 ± 0.212	0.376 ± 0.18	0.879 ± 0.088	0.628 ± 0.096
22	0.736 ± 0.077	0.255 ± 0.25	0.35 ± 0.222	0.874 ± 0.087	0.612 ± 0.11
23	**0.785 ± 0.077**	**0.431 ± 0.212**	**0.502 ± 0.187**	0.886 ± 0.088	**0.694 ± 0.099**
24	0.771 ± 0.081	0.392 ± 0.22	0.488 ± 0.203	0.872 ± 0.094	0.68 ± 0.105
25	0.776 ± 0.088	0.398 ± 0.247	0.488 ± 0.225	0.879 ± 0.094	0.683 ± 0.119
26	0.774 ± 0.087	0.388 ± 0.247	0.464 ± 0.204	0.885 ± 0.095	0.675 ± 0.111
27	0.762 ± 0.081	0.351 ± 0.229	0.428 ± 0.202	0.881 ± 0.093	0.654 ± 0.106
28	0.768 ± 0.077	0.352 ± 0.24	0.402 ± 0.193	0.899 ± 0.082	0.65 ± 0.103
29	0.782 ± 0.078	0.405 ± 0.233	0.462 ± 0.222	0.896 ± 0.089	0.679 ± 0.11
30	**0.785 ± 0.08**	0.422 ± 0.227	0.5 ± 0.213	0.886 ± 0.085	0.693 ± 0.11

Table 21. The *RFECV* features performance measures for the *Full* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
19	**0.743 ± 0.079**	**0.283 ± 0.227**	**0.356 ± 0.187**	**0.881 ± 0.095**	**0.619 ± 0.098**
21	0.719 ± 0.08	0.232 ± 0.202	**0.356 ± 0.169**	0.849 ± 0.092	0.602 ± 0.095

The first feature set was achieved by using the *accuracy* and the second one with the *average precision* metric.

Table 22. The *UFS+RFE* features performance measures for the *Full* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.858 ± 0.069	0.629 ± 0.173	0.63 ± 0.18	**0.939 ± 0.076**	0.785 ± 0.093
3	0.796 ± 0.107	0.54 ± 0.23	0.724 ± 0.185	0.821 ± 0.135	0.773 ± 0.11
4	0.839 ± 0.09	0.612 ± 0.207	0.724 ± 0.189	0.881 ± 0.102	0.802 ± 0.107
5	0.842 ± 0.075	0.625 ± 0.185	0.766 ± 0.208	0.869 ± 0.093	0.817 ± 0.101
6	0.834 ± 0.074	0.608 ± 0.183	0.776 ± 0.194	0.854 ± 0.081	0.815 ± 0.101
7	0.821 ± 0.085	0.57 ± 0.204	0.71 ± 0.218	0.861 ± 0.105	0.785 ± 0.109
8	0.829 ± 0.082	0.608 ± 0.188	**0.778 ± 0.19**	0.847 ± 0.103	0.813 ± 0.099
9	0.838 ± 0.068	0.622 ± 0.157	0.776 ± 0.188	0.86 ± 0.094	0.818 ± 0.088
10	0.844 ± 0.072	0.613 ± 0.198	0.724 ± 0.226	0.887 ± 0.082	0.806 ± 0.109
11	0.836 ± 0.075	0.604 ± 0.184	0.732 ± 0.192	0.873 ± 0.095	0.802 ± 0.096
12	0.842 ± 0.075	0.607 ± 0.191	0.712 ± 0.201	0.889 ± 0.085	0.8 ± 0.103
13	0.872 ± 0.07	0.684 ± 0.183	0.766 ± 0.21	0.91 ± 0.078	**0.838 ± 0.104**
14	0.862 ± 0.07	0.649 ± 0.189	0.718 ± 0.214	0.914 ± 0.072	0.816 ± 0.106
15	**0.873 ± 0.068**	**0.686 ± 0.174**	0.764 ± 0.188	0.912 ± 0.075	**0.838 ± 0.096**
16	0.869 ± 0.076	0.665 ± 0.213	0.72 ± 0.232	0.923 ± 0.077	0.821 ± 0.116
17	0.86 ± 0.063	0.65 ± 0.155	0.71 ± 0.193	0.914 ± 0.08	0.812 ± 0.091
18	0.848 ± 0.077	0.616 ± 0.21	0.702 ± 0.218	0.901 ± 0.078	0.801 ± 0.112
19	0.858 ± 0.07	0.639 ± 0.187	0.708 ± 0.214	0.911 ± 0.08	0.81 ± 0.104
20	0.863 ± 0.079	0.647 ± 0.215	0.714 ± 0.212	0.916 ± 0.072	0.815 ± 0.114
21	0.872 ± 0.066	0.672 ± 0.185	0.728 ± 0.218	0.923 ± 0.07	0.825 ± 0.105
22	0.857 ± 0.062	0.643 ± 0.159	0.724 ± 0.176	0.905 ± 0.071	0.815 ± 0.087
23	0.85 ± 0.07	0.619 ± 0.18	0.68 ± 0.19	0.911 ± 0.08	0.795 ± 0.096
24	0.863 ± 0.065	0.649 ± 0.177	0.704 ± 0.184	0.92 ± 0.069	0.812 ± 0.094
25	0.853 ± 0.062	0.628 ± 0.162	0.698 ± 0.191	0.908 ± 0.076	0.803 ± 0.091
26	0.86 ± 0.065	0.637 ± 0.19	0.688 ± 0.21	0.921 ± 0.067	0.805 ± 0.103
27	0.865 ± 0.059	0.651 ± 0.162	0.69 ± 0.199	0.927 ± 0.065	0.809 ± 0.095
28	0.866 ± 0.067	0.66 ± 0.177	0.712 ± 0.197	0.921 ± 0.076	0.816 ± 0.097
29	0.866 ± 0.069	0.666 ± 0.18	0.75 ± 0.188	0.907 ± 0.077	0.829 ± 0.096
30	0.864 ± 0.065	0.656 ± 0.179	0.724 ± 0.205	0.914 ± 0.075	0.819 ± 0.099

Table 23. The *RF* performance measures for the *Full* dataset.

F	Accuracy	MCC	Sensitivity	Specificity	AUC
2	0.837 ± 0.063	0.55 ± 0.203	0.522 ± 0.215	**0.949 ± 0.061**	0.736 ± 0.104
3	0.839 ± 0.074	0.571 ± 0.207	0.596 ± 0.194	0.926 ± 0.071	0.761 ± 0.104
4	0.835 ± 0.062	0.561 ± 0.172	0.57 ± 0.186	0.93 ± 0.069	0.75 ± 0.091
5	0.852 ± 0.059	0.618 ± 0.163	0.678 ± 0.185	0.914 ± 0.066	0.796 ± 0.09
6	0.846 ± 0.082	0.634 ± 0.199	**0.776 ± 0.182**	0.871 ± 0.088	0.824 ± 0.103
7	0.858 ± 0.062	0.648 ± 0.157	0.744 ± 0.179	0.899 ± 0.072	0.821 ± 0.088
8	0.857 ± 0.069	0.647 ± 0.179	0.75 ± 0.199	0.895 ± 0.077	0.823 ± 0.099
9	0.842 ± 0.067	0.606 ± 0.17	0.716 ± 0.177	0.887 ± 0.073	0.802 ± 0.091
10	0.846 ± 0.066	0.607 ± 0.185	0.69 ± 0.21	0.902 ± 0.073	0.796 ± 0.101
11	0.838 ± 0.067	0.596 ± 0.172	0.694 ± 0.189	0.89 ± 0.079	0.792 ± 0.093
12	0.849 ± 0.078	0.627 ± 0.192	0.72 ± 0.188	0.896 ± 0.085	0.808 ± 0.101
13	0.856 ± 0.073	0.641 ± 0.186	0.738 ± 0.189	0.898 ± 0.074	0.818 ± 0.1
14	**0.87 ± 0.072**	**0.683 ± 0.169**	0.77 ± 0.175	0.906 ± 0.084	**0.838 ± 0.092**
15	0.855 ± 0.082	0.645 ± 0.201	0.74 ± 0.203	0.896 ± 0.091	0.818 ± 0.108
16	0.853 ± 0.068	0.628 ± 0.179	0.694 ± 0.193	0.91 ± 0.078	0.802 ± 0.096
17	0.855 ± 0.064	0.635 ± 0.164	0.712 ± 0.17	0.906 ± 0.073	0.809 ± 0.087
18	0.852 ± 0.075	0.636 ± 0.191	0.74 ± 0.191	0.892 ± 0.082	0.816 ± 0.1
19	0.861 ± 0.081	0.657 ± 0.205	0.758 ± 0.19	0.898 ± 0.083	0.828 ± 0.105
20	0.857 ± 0.063	0.649 ± 0.16	0.746 ± 0.192	0.896 ± 0.078	0.821 ± 0.091
21	0.842 ± 0.079	0.628 ± 0.184	0.77 ± 0.177	0.868 ± 0.093	0.819 ± 0.096
22	0.848 ± 0.067	0.629 ± 0.169	0.744 ± 0.188	0.886 ± 0.079	0.815 ± 0.093
23	0.843 ± 0.07	0.621 ± 0.176	0.744 ± 0.196	0.879 ± 0.089	0.811 ± 0.094
24	0.854 ± 0.066	0.639 ± 0.166	0.722 ± 0.196	0.901 ± 0.08	0.812 ± 0.094
25	0.856 ± 0.067	0.643 ± 0.175	0.742 ± 0.18	0.897 ± 0.074	0.82 ± 0.092
26	0.852 ± 0.063	0.632 ± 0.146	0.706 ± 0.154	0.904 ± 0.082	0.805 ± 0.077
27	0.846 ± 0.069	0.628 ± 0.18	0.752 ± 0.186	0.88 ± 0.086	0.816 ± 0.092
28	0.839 ± 0.069	0.612 ± 0.162	0.738 ± 0.189	0.875 ± 0.091	0.806 ± 0.09
29	0.845 ± 0.072	0.624 ± 0.175	0.744 ± 0.186	0.881 ± 0.084	0.813 ± 0.095
30	0.832 ± 0.062	0.587 ± 0.154	0.708 ± 0.166	0.876 ± 0.079	0.792 ± 0.081

A.3.2 Feature sets rankings

Table 24. The *UFS* and *RFE* feature rankings for the *Full* dataset.

R	Univariate feature selection	Recursive feature elimination
1	APACHE II_T1	A wave cm s _T2
2	Creatinine mg dL _T3	APACHE II_T1
3	E e _T3	CRP_Value mg L _T3
4	FiO2_T2	Creat Ur_T3
5	FiO2_T3	Diastolic Blood Pressure mmHg _T3
6	Glasgow Coma Scale_T2	E wave cm s _T1
7	Heart rate bpm _T3 1	HCO3 mmol L _T1
8	Hypotension SBP 90 mmHg or MAP 70 mmHg or SBP decrease 40 mmHg _T2=No	Heart rate bpm _T2 1
9	Hypotension SBP 90 mmHg or MAP 70 mmHg or SBP decrease 40 mmHg _T2=Yes	Heart rate bpm _T3
10	Inferior vena cava distensibility index _T3	Hematocrit _T1
11	LA dilatation by eyeballing_T1=No	K Ur_T1
12	LA dilatation by eyeballing_T1=Yes	K Ur_T3
13	Lactate levels mmol L _T2	Leukocytes total 1000 mm 3 _T2
14	Na Ur_T3	Neutro abs count_T2
15	Number of affected organs	PCT_Value mg mL _T2
16	PCT_Value mg mL _T2	PT_T1
17	RBC count_T3	PaCO2 mmHg _T1
18	Respiratory rate_T1	Platelet count_T1
19	SOFA_T1	Pplat_T3
20	SOFA_T2	Propofol μg kg min _T1
21	SOFA_T3	Respiratory rate rpm _T1
22	Sat O2 FiO2_T2	Respiratory rate_T2
23	Sat O2 FiO2_T3	SvO2 _T2
24	Sedation Scale SAS _T2	Tidal volume VT _T1
25	Tracheal Intubation_T2=No	Tidal volume VT _T3
26	Tracheal Intubation_T2=Yes	Tricuspid regurgitation maximal velocity by CW cm s _T1
27	Tricuspid regurgitation maximal velocity by CW cm s _T1	Tricuspid regurgitation maximal velocity by CW cm s _T2
28	X Norepinephrine mg kg min _T1	Value _T1
29	X Norepinephrine mg kg min _T2	WBC abs count_T2
30	pH_T2	aPTT_T1

```
ACCURACY: Creat Ur_T3, Diastolic Blood Pressure mmHg _T3, E wave cm s
    _T1, HCO3 mmol L _T1, Heart rate bpm _T3, Hematocrit _T1, K Ur_T1,
    Leukocytes total 1000 mm 3 _T2, Neutro abs count_T2, PT_T1, Platelet
     count_T1, Propofol mg kg min _T1, Tidal volume VT _T1, Tidal volume
     VT _T3, Tricuspid regurgitation maximal velocity by CW cm s _T1,
    Tricuspid regurgitation maximal velocity by CW cm s _T2, Value _T1,
    WBC abs count_T2, aPTT_T1;
```

```
AVERAGE PRECISION: Creat Ur_T3, Diastolic Blood Pressure mmHg _T3, E
    wave cm s _T1, HCO3 mmol L _T1, Heart rate bpm _T3, Hematocrit _T1,
    K Ur_T1, K Ur_T3, Leukocytes total 1000 mm 3 _T2, Neutro abs
    count_T2, PCT_Value mg mL _T2, PT_T1, Platelet count_T1, Propofol mg
     kg min _T1, Tidal volume VT _T1, Tidal volume VT _T3, Tricuspid
    regurgitation maximal velocity by CW cm s _T1, Tricuspid
    regurgitation maximal velocity by CW cm s _T2, Value _T1, WBC abs
    count_T2, aPTT_T1.
```

List of features for the *RFECV* with the *accuracy* and the *average precision* metrics for the **Full** dataset.

Table 25. The *UFS+RFE* and *RF* feature rankings for the *Full* dataset.

R	UFS+RFE	Random Forest
1	Inferior vena cava distensibility index _T3	Inferior vena cava distensibility index _T3
2	LA dilatation by eyeballing_T1=No	Lactate levels mmol L _T2
3	LA dilatation by eyeballing_T1=Yes	SOFA_T2
4	Is the patient under other drugs_T1=No	SOFA_T3
5	Glasgow Coma Scale_T3	Na Ur_T3
6	Is the patient under other drugs_T1=Yes	APACHE II_T1
7	Number of affected organs	FiO2_T3
8	SOFA_T3	PCT_Value mg mL _T1
9	Pulmonary artery systolic pressure TR jet by CW CVP mmHg _T2	Tricuspid regurgitation maximal velocity by CW cm s _T3
10	Glasgow Coma Scale_T1	X Norepinephrine mg kg min _T2
11	Lactate levels mmol L _T1	Respiratory rate_T1
12	SOFA_T2	Urine Output mL day _T1
13	PCT_Value mg mL _T3	Sat O2 FiO2_T3
14	Hematocrit _T3	SOFA_T1
15	APACHE II_T1	Tidal volume VT _T2
16	Heart rate bpm _T3 1	X Dobutamine mg kg min _T3
17	K Ur_T1	K Ur_T1
18	Respiratory rate_T1	E wave cm s _T1
19	Tricuspid regurgitation maximal velocity by CW cm s _T1	Glasgow Coma Scale_T3
20	LVOT average diameter mm _T1	RBC count_T3
21	CRP_Value mg L _T3	Na Ur_T2
22	PCT_Value mg mL _T1	PCT_Value mg mL _T2
23	X Dobutamine mg kg min _T3	Urine Output mL day _T2
24	SOFA_T1	Lateral e cm s _T3
25	Glasgow Coma Scale_T2	Tricuspid annular plane systolic excursion TAPSE mm _T3
26	X Norepinephrine mg kg min _T1	E e _T3
27	Platelet count_T1	K Ur_T2
28	Creatinine mg dL _T3	Fluid Balance ml _T2
29	Base Excess mmol L _T1	X Norepinephrine mg kg min _T1
30	Platelets 10 3 mm 3 _T3	SvcO2 _T2

A.4 Stability scores

Table 26. *T1* stability scores.

T1		Random Forest	
LA dilatation by eyeballing_-T1=Yes	0.778	APACHE II_T1	0.898
E A ratio_T1=0.6	0.778	Urine Output mL day _T1	0.882
Sedation Scale SAS _T1	0.778	Respiratory rate_T1	0.878
Is the patient under other drugs_-T1=No	0.778	E wave cm s _T1	0.844
Glasgow Coma Scale_T1	0.778	Lactate levels mmol L _T1	0.822
Is the patient under other drugs_-T1=Yes	0.778	Tidal volume VT _T1	0.808
SOFA_T1	0.667	SOFA_T1	0.804
Respiratory rate_T1	0.667	Fluid Balance ml _T1	0.794
Aortic Valve Regurgitation_-T1=Moderate	0.667	E e _T1	0.786
Aortic Valve Regurg._T1=No val.dysfunc.	0.667	X Norepinephrine mg kg min _T1	0.782
HCO3 mmol L _T1	0.667	Heart rate bpm _T1	0.776
LA dilatation by eyeballing_-T1=No	0.667	Pplat_T1	0.764
Number of affected organs	0.556	Temperature C _T1	0.72
Base Excess mmol L _T1	0.556	pH_T1	0.706
LV Hypertrophy LV mass _-T1=Moderate	0.556	Fibrinogen_T1 g L	0.706
Weight kg	0.444	Respiratory rate rpm _T1	0.7
LV Hypertrophy LV mass _-T1=No	0.444	Platelet count_T1	0.68
Sat O2 _T1	0.444	PEEP_T1	0.676
Lateral e cm s _T1	0.444	Leukocytes total 1000 mm 3 _T1	0.656
aPTT_T1	0.444	LVOT avg.vel.time integral_T1	0.656
Midazolam mg kg min _T1	0.444	PaO2 FiO2_T1	0.654
Fibrinogen_T1 g L	0.444	Midazolam mg kg min _T1	0.65
Respiratory rate rpm _T1	0.444	Base Excess mmol L _T1	0.64
LVOT avg. vel. time integral cm _T1	0.444	Neutro abs count_T1	0.64
PT_T1	0.444	A wave cm s _T1	0.632
Heart rate bpm _T1 1	0.444	Glycemia mg dL _T1	0.628
Neutro abs count_T1	0.444	Platelets 10 3 mm 3 _T1	0.628
Diagnosed within 48 from admission=No	0.444	RBC count_T1	0.622
Hematocrit _T1	0.444	LVOT average diameter mm _T1	0.62
Pplat_T1	0.444	Lympho abs count_T1	0.616

The first 30 features stability scores using the **T1** dataset: for the *RF* (Random Forest) feature selection and for the rest of the feature selection techniques (*T1*)

Table 27. *T1+T2* stability scores.

T1+T2		Random Forest	
Glasgow Coma Scale_T1	1.0	SOFA_T2	0.766
Base Excess mmol L _T1	0.8	Lactate levels mmol L _T2	0.74
SOFA_T2	0.8	APACHE II_T1	0.722
SOFA_T1	0.8	Respiratory rate_T1	0.704
HCO3 mmol L _T1	0.6	Urine Output mL day _T1	0.7
LVOT average diameter mm _T1	0.6	Urine Output mL day _T2	0.698
APACHE II_T1	0.6	X Norepinephrine μg kg min _T1	0.686
Glasgow Coma Scale_T2	0.6	E wave cm s _T1	0.638
Platelet count_T1	0.6	pH_T2	0.628
Tidal volume VT _T1	0.6	Pplat_T2	0.624
Respiratory rate_T2	0.4	SOFA_T1	0.622
Respiratory rate_T1	0.4	Tidal volume VT _T1	0.622
pH_T2	0.4	Fluid Balance ml _T2	0.62
Weight kg	0.4	Glasgow Coma Scale_T2	0.604
A wave cm s _T1	0.4	Creatinine mg dL _T2	0.6
Sedation Scale SAS _T2	0.4	Lactate levels mmol L _T1	0.592
Na mmol L _T1	0.4	Fluid Balance ml _T1	0.576
Hematocrit _T1	0.4	Respiratory rate_T2	0.558
Heart rate bpm _T1 1	0.4	Heart rate bpm _T1	0.558
Hematocrit _T2	0.4	E e _T1	0.558
Anuria and or RRT_T1=Yes	0.4	PT_T2	0.544
Aortic Valve Stenosis_T1=Moderate	0.4	FiO2_T2	0.542
E A ratio_T1=1.39	0.4	Platelet count_T2	0.528
Pplat_T2	0.4	Sat O2 _T2	0.524
Number of affected organs	0.4	Lympho abs count_T2	0.518
aPTT_T2	0.4	PaO2 mmHg _T2	0.516
LV Dilatation LVEDV _T1=Moderate	0.4	PEEP_T2	0.51
PT_T2	0.4	Temperature C _T1	0.496
PT_T1	0.4	pH_T1	0.496
SvcO2 _T1	0.4	Pplat_T1	0.494

The first 30 features stability scores using the **T1+T2** dataset: for the *RF* (Random Forest) feature selection and for the rest of the feature selection techniques (*T1+T2*).

Table 28. *Full* stability scores.

Full		Random Forest	
Tricuspid regurgitation maximal velocity by CW cm s _T1	1.0	Inferior vena cava distensibility index _T3	0.728
K Ur_T1	0.8	Lactate levels mmol L _T2	0.612
Platelet count_T1	0.8	SOFA_T2	0.582
Creat Ur_T3	0.6	SOFA_T3	0.542
Neutro abs count_T2	0.6	Na Ur_T3	0.504
HCO3 mmol L _T1	0.6	APACHE II_T1	0.5
Hematocrit _T1	0.6	FiO2_T3	0.476
Value _T1	0.6	PCT_Value mg mL _T1	0.464
Diastolic Blood Pressure mmHg _T3	0.6	Tricuspid regurgitation maximal velocity by CW cm s _T3	0.464
APACHE II_T1	0.6	X Norepinephrine mg kg min _T2	0.454
Propofol μg kg min _T1	0.6	Respiratory rate_T1	0.452
PCT_Value mg mL _T2	0.6	Urine Output mL day _T1	0.446
PT_T1	0.6	Sat O2 FiO2_T3	0.43
E wave cm s _T1	0.6	SOFA_T1	0.422
Leukocytes total 1000 mm 3 _T2	0.6	Tidal volume VT _T2	0.422
Tricuspid regurgitation maximal velocity by CW cm s _T2	0.6	X Dobutamine mg kg min _T3	0.408
Tidal volume VT _T3	0.6	K Ur_T1	0.402
aPTT_T1	0.6	E wave cm s _T1	0.4
Heart rate bpm _T3	0.6	Glasgow Coma Scale_T3	0.4
Tidal volume VT _T1	0.6	RBC count_T3	0.394
WBC abs count_T2	0.6	Na Ur_T2	0.386
Respiratory rate_T1	0.4	PCT_Value mg mL _T2	0.38
Creatinine mg dL _T3	0.4	Urine Output mL day _T2	0.38
CRP_Value mg L _T3	0.4	Lateral e cm s _T3	0.374
LA dilatation by eyeballing_T1=Yes	0.4	Tricuspid annular plane systolic excursion TAPSE mm _T3	0.372
X Norepinephrine mg kg min _T1	0.4	E e _T3	0.37
K Ur_T3	0.4	K Ur_T2	0.364
Heart rate bpm _T3 1	0.4	Fluid Balance ml _T2	0.36
LA dilatation by eyeballing_T1=No	0.4	X Norepinephrine mg kg min _T1	0.356
Inferior vena cava distensibility index _T3	0.4	SvcO2 _T2	0.354

The first 30 features stability scores using the **Full** dataset: for the *RF* (Random Forest) feature selection and for the rest of the feature selection techniques (*Full*).