

# Angiotensin-1, Angiotensin-2 and Bicarbonate as Diagnostic Biomarkers in Children with Severe Sepsis

Kun Wang<sup>1,2</sup>, Vineet Bhandari<sup>3</sup>, John S. Giuliano Jr<sup>3</sup>, Corey S. O'Hern<sup>2,4</sup>, Mark D. Shattuck<sup>5</sup>, Michael Kirby<sup>1,\*</sup>

**1 Department of Mathematics, Colorado State University, Fort Collins, CO, USA**

**2 Department of Mechanical Engineering & Materials Science, Yale University, New Haven, CT, USA**

**3 Department of Pediatrics, Yale University School of Medicine, New Haven, CT, USA**

**4 Department of Applied Physics and Department of Physics, Yale University, New Haven, CT, USA**

**5 Benjamin Levich Institute and Physics Department, The City College of New York, New York, NY, USA**

**\* E-mail: kirby@math.colostate.edu**

## Abstract

We employ machine learning techniques to validate the use of Angiotensin-1 (Ang-1), Ang-2 and Bicarbonate ( $\text{HCO}_3$ ) levels as biomarkers for determining the severity of sepsis in critically ill children. The study was based on clinical data and plasma samples provided by Yale-New Haven Children's Hospital's Pediatric Intensive Care Unit (PICU) from a group of 45 patients with varying sepsis severity at the time of admission. Canonical Correlation Analysis with the Forward Selection and Random Forests methods identified the above biomarkers as having the strongest correlations with sepsis severity. The robustness and effectiveness of these biomarkers were validated by constructing a linear Support Vector Machine diagnostic classifier. Our study demonstrates a consistent pattern for patients with septic shock characterized by 1) decreased levels of Ang-1, which stabilizes capillaries and promotes endothelial cell survival, 2) increased levels of Ang-2, which destabilizes endothelial cell junctions and promotes inflammatory cell migration, and 3) decreased levels of  $\text{HCO}_3$ , which is an indicator of an acid-base imbalance in the body. We also show that the concentrations of Ang-1, Ang-2 and  $\text{HCO}_3$  enable predictions of the time dependence of sepsis severity in children.

## Introduction

Diagnosing and classifying the severity of sepsis is a significant challenge due to the highly variable and nonspecific nature of the signs and symptoms of sepsis. Biomarkers which play critical roles in the disease process, show great promise in indicating the severity of

sepsis. There are many biomarkers that have been studied for potential use in the early diagnosis and classification of sepsis [1, 2].

Among novel biomarkers, studies have shown that Angiotensin-1 (Ang-1) and Ang-2 are among the most promising endothelial associated sepsis biomarkers in several populations [3–6]. In particular, levels of Ang-1 and Ang-2 have been shown to be associated with the severity of sepsis in children [5, 7]. These studies demonstrated a consistent pattern in children with severe sepsis, characterized by increased levels of Ang-2 and decreased levels of Ang-1. Others have shown that Ang-2 disrupts endothelial cell barrier function, while Ang-1 ameliorates the same, in sepsis [8]. Therefore, the relative concentration of Ang-2 and Ang-1, *i.e.* the Ang-2/Ang1 ratio, has been identified as a more robust biomarker than isolated levels of each factor [5, 7]. In addition, the serum concentration of bicarbonate ( $\text{HCO}_3$ ) is an important measure of the acid-base balance during sepsis [9].

However, none of these biomarkers taken individually has sufficient specificity or sensitivity to be routinely employed in clinical practice. A combination of several sepsis biomarkers may be more effective, as has been suggested by other investigators [10–12]. In this manuscript, we present a systematic study of the multivariate diagnostic capacity of a scoring system that incorporates Ang-1, Ang-2 and  $\text{HCO}_3$  to distinguish patients admitted to Yale-New Haven Children’s Hospital’s Pediatric Intensive Care Unit (PICU) with or without sepsis (PICU/sepsis group) versus those with severe sepsis (PICU severe sepsis group).

## Materials

### Study population

This study was approved by the Pediatric Protocol Review Committee and the Human Investigation Committee at Yale University School of Medicine. The biological specimens and clinical data sets were obtained from a prospective observational study of critically ill pediatric patients with varying degrees of sepsis severity conducted at a tertiary care center PICU during the time period 9/2009-12/2011 [5].

All patients admitted to the PICU were evaluated for eligibility. Forty-five patients met the eligibility criteria and consented to participate in the study. Blood samples were collected every 12 hours for the first 3 days and then once a day for the last 4 days. Data collection was discontinued when the patient was discharged from the PICU. A maximum of 10 samples for 7 days were obtained from each patient. Commercial enzyme-linked immunosorbent assay (ELISA) kits were used to measure plasma levels of Ang-1 and Ang-2. Descriptive data consisting of demographics and clinical data were also recorded for all patients. Additional details have been recently published [5].

## Biomarkers

To create a robust model of a specific combination of biomarkers for predicting the severity of sepsis in children in an unbiased manner, we selected multiple clinical and laboratory variables from the database of our study [5]. These 17 variables are as follows: (1) Age, (2) Weight (Wgt), (3) admission Pediatric Index of Mortality 2 (PIM-2) [13], (4) White Blood Cell count (WBC), (5) Hemoglobin count (Hgb), (6) Hematocrit (Hct), (7) Platelet count (Plt), and the levels of (8) Sodium (Na), (9) Potassium (K), (10) Chloride (Cl), (11)  $\text{HCO}_3$ , (12) Blood Urea Nitrogen (BUN), (13) Creatinine (Cr), (14) Ang-1, (15) Ang-2, (16) Ang-2/Ang-1 ratio, and (17) Vascular Endothelial Growth Factor (VEGF). To validate the data analysis, we augmented this data set to include (18) Gaussian distributed noise (g-Noise) and (19) uniformly distributed noise (u-Noise). These 19 variables were then used to develop sepsis severity prediction models.

## Statistical Analysis

Patients were classified within the first 24 hours of PICU admission into the following two categories based on the 2005 pediatric sepsis and organ dysfunction definitions [14]: 1) the PICU/sepsis group ( $n = 28$ ) included those with nonsystemic inflammatory response syndrome (non-SIRS) ( $n = 9$ ), SIRS ( $n = 8$ ), and sepsis ( $n = 11$ ), and 2) the PICU severe sepsis group ( $n = 17$ ) included those with severe sepsis ( $n = 3$ ), and septic shock ( $n = 14$ ). We systematically evaluated the Pearson pairwise correlations between all pairs of biomarkers for the PICU/sepsis and PICU severe sepsis groups.

## Results

### Biomarkers Selection

Feature selection is an important part of the data analysis given the fact that the data contains many redundant or irrelevant features. Redundant features provide no additional information than the selected features, and irrelevant features provide no useful information. Feature selection is widely used in data sets with abundant features but comparatively few samples. In machine learning and statistics, the goal of a feature selection method is to select an optimal subset of relevant features for model construction.

In this study, there are 17 variables (features) augmented by 2 variables consisting of Gaussian and uniform noise to provide a baseline check for the data analysis. From the univariate correlation analysis, we found that this data set contained several possible redundant biomarkers and, not surprisingly, at least two irrelevant features (g-Noise and u-Noise). To extract an optimal subset of biomarkers, we analyzed the multivariate correlation between the outcome, sepsis severity score (0 for PICU/sepsis and 1 for PICU severe sepsis), and the input, which is a subset of variables.

A comparison of the univariate correlations for these two groups is shown in Fig. 1. The univariate analysis revealed that Na, K, Cl,  $\text{HCO}_3$  form a group of highly correlated biomarkers (with correlations that range from 0.937 to 0.998) for the PICU/sepsis group. However, these variables are not strongly correlated for the PICU severe sepsis group (with correlations that range from 0.001 to 0.608). This notable difference between the PICU/sepsis and PICU severe sepsis groups indicates that these biomarkers may not independently provide information about the sepsis severity diagnosis. We also note that Ang-1 and Ang-2 are highly correlated with each other in the PICU severe sepsis group (0.76), but this correlation is significantly reduced for the PICU/sepsis group (0.21). Meanwhile, Ang-2/Ang-1 does not correlate very strongly with either Ang-1 (0.21 in PICU/sepsis, 0.24 in PICU severe sepsis) or Ang-2 (0.48 in PICU/sepsis, 0.17 in PICU severe sepsis). Based on these observations, we seek to identify an optimal set of non-redundant variables and biomarkers to predict the severity of sepsis.

In our recent study [15], we found that canonical correlation analysis (CCA) [16–18] can be applied effectively to identify an optimal subset of biomarkers with the maximum correlation with the outcome. As shown in Table 1, we found that the subset of Ang-2, Ang-1, and  $\text{HCO}_3$  maximizes the correlation with the sepsis severity score. As expected, the two forms of random noise are selected near the end of the process when the correlation saturates for large subsets. We also applied the forward selection (FS) method to identify the optimal subset of biomarkers. FS is a greedy algorithm that adds the best feature at each step [19,20]. We found that the performance of the subset of biomarkers selected by FS was similar to that selected by CCA on this data set.

## The Optimal Subset

In this study, we built a diagnostic classifier by selecting the subset of  $k$  biomarkers with the best diagnostic performance for each value of  $k$ . For each  $k$ , we applied the ensemble method [21,22] to construct a linear support vector machine (SVM) classifier [23] for the CCA-selected subset of biomarkers. SVM [24] finds a decision function that separates the high-dimensional data with the maximum margin. To quantify the classifier performance, we calculated the true positive rate (TPR), true negative rate (TNR), positive predictive value (PPV), and negative predictive value (NPV). See the Methods section for details.

In Figure 2, we find that all statistical measures reach a peak or saturate near  $k = 3$  using the CCA-selected biomarkers, Ang-2, Ang-1, and  $\text{HCO}_3$ , which suggests that these three biomarkers are the optimal subset for our data set (TPR = 0.69, TNR = 0.87, PPV = 0.79, and NPV = 0.83 at  $k = 3$ ). By adding  $\text{HCO}_3$  to the optimal subset from  $k = 2$  to  $k = 3$ , the combination has higher TPR (0.60 at  $k = 2$  versus 0.69 at  $k = 3$ ) and PPV (0.69 at  $k = 2$  versus 0.79 at  $k = 3$ ) when compared to the combination of Ang-2 and Ang-1. TNR (0.84 at  $k = 6$  versus 0.80 at  $k = 7$ ) and PPV (0.75 at  $k = 6$  versus 0.69 at  $k = 7$ ) begin to decrease from their plateau values when  $\text{HCO}_3$  leaves the subset at  $k = 7$ . The improvement at  $k = 3$  and decrease at  $k = 7$  indicate the diagnostic

importance of  $\text{HCO}_3$ .

## Redundant Biomarkers

Recent studies [3–5,7] suggest that plasma levels of Ang-2 and Ang-1 can serve as clinically informative biomarkers of sepsis severity. Further, the Ang-2/Ang1 ratio is considered to be a more relevant sepsis severity biomarker than isolated levels of each biomarker because of their antagonistic roles in regulating the tyrosine kinase receptor, Tie-2 [7]. However, both of our biomarker selection methods, CCA and FS, select Ang-2/Ang-1 to the optimal subset relatively late, *i.e.*, at large  $k$  ( $k = 13$  and  $k = 15$ ) as shown in Table 1. This suggests that a combination of Ang-2, Ang-1, and  $\text{HCO}_3$ , is potentially more effective than using the ratio of Ang-1 and Ang-2 with other biomarkers.

It is also interesting to consider the univariate and bivariate performance of these biomarkers. This analysis provides additional insight into the relative performance of different subsets of biomarkers and how they work together to provide inferences.

In Fig. 3(A), the relative performance of the univariate biomarkers performance is shown: 1) Ang-1 has consistent performance for all statistical measures compared to other biomarkers (see Table 2), 2) Ang-2 has a high TNR (0.85) and PPV (0.63) but relatively low TPR (0.38), and 3)  $\text{HCO}_3$  has the highest TPR (0.87) and NPV (0.86) but relatively low TNR (0.42) and PPV (0.48). These observations indicate that the performances of these biomarkers did not correlate with each other. This supports the observation that the best subset of biomarkers includes both Ang-1 and Ang-2 since they provide distinct information. We also show that the combination of Ang-2, Ang-1 and  $\text{HCO}_3$  improves the predictive capability by reducing overfitting in Fig. 2. The performance for the CCA-selected subsets decreases relatively when  $k > 3$ .

These results suggest, when examining groups of three, Ang-2/Ang-1 may be a redundant biomarker, *i.e.*, no additional information is gained when Ang-1 and Ang-2 data is known. We explore here how this ratio performs in isolation, *i.e.*, as a derived univariate statistic. We applied the same procedure as above to construct a SVM classifier for each single biomarker and show the statistical measures in Fig. 3(A). Overall, we find that Ang-2 and Ang-2/Ang-1 have comparable prediction performance (Fig. 3(A)). However, Ang-2/Ang-1 outperforms Ang-2 for PPV (0.76 for Ang-2/Ang-1, 0.63 for Ang-2), which suggests that Ang-2/Ang-1 alone may be a predictive biomarker. The similar performance of Ang-2 and Ang-2/Ang-1 suggest that these two biomarkers capture very similar information.

Of course it is not necessarily a fair assessment to compare true univariate biomarkers such as Ang-1 and Ang-2 to their ratio since this contains information from two measurements. Thus, we also compared the performance of combinations of Ang-1, Ang-2,  $\text{HCO}_3$ , and Ang-2/Ang-1 in Fig. 3(B). The combination of Ang-2 and Ang-2/Ang-1 does not notably improve each predictive measure compared to these biomarkers alone, which also indicates that these two biomarkers are redundant. In contrast, the combination of

Ang-1 and Ang-2 has notably higher NPV (0.78) and TPR (0.60) and comparable values for the other prediction measures compared to each single biomarker (NPV = 0.69 and TPR = 0.38 for Ang-2, NPV = 0.57 and TPR = 0.48 for Ang-1) and Ang-2/Ang-1 (NPV = 0.67 and TPR = 0.35). This suggests that the ratio Ang-2/Ang-1 is less effective than using Ang-1 and Ang-2 separately.

For completeness, we also show the performance for the CCA-selected optimal subset of three biomarkers  $\text{HCO}_3$ , Ang-1 and Ang-2 on the far right of Fig. 3(B). This optimal subset notably improves the predictive capability as indicated by the small spread of values in the predictive measures.

## The Diagnostic Classifier

We applied the linear SVM ensemble method [22, 23] to construct a decision function using the CCA-selected optimal subset of biomarkers at  $k = 3$ : Ang-2, Ang-1, and  $\text{HCO}_3$ . The optimal decision function is

$$\text{Score} = w_1\text{Ang-2} + w_2\text{Ang-1} + w_3\text{HCO}_3 - b. \quad (1)$$

Table 3 provides the weights  $w_i$ , errors  $e_i$ , means  $\bar{x}_i$  and standard deviations  $\sigma_i$  of the biomarkers. Since the range of values of the biomarkers varies widely, all values of the biomarkers are normalized by subtracting the mean and then dividing by the standard deviation in Eq. 1. See Methods for details. With this decision function, if the sepsis severity score (Score) is greater than or equal to zero, the severity diagnosis is 1, otherwise it is 0. The magnitudes of weights  $w_i$  indicate the importance of the corresponding biomarker [25]. We find that Ang-2 has a larger weight than Ang-1 and  $\text{HCO}_3$ , which is consistent with the results for the single biomarker classification in Fig. 3(A), where the TNR, and PPV are larger for Ang-2 than Ang-1 and  $\text{HCO}_3$ . However, the TPR and NPV are larger for  $\text{HCO}_3$  compared to that for Ang-2. The signs of weights  $w_i$  indicates the sign of the correlation of the biomarker with the sepsis severity score. Thus, the sepsis severity score for a patient with a relatively high Ang-2 level and low Ang-1 and  $\text{HCO}_3$  levels is most likely positive. This relation between biomarkers and sepsis severity score has been observed in the clinical studies [3, 26, 27].

## Longitudinal Measurements of the Predictor

A linear SVM finds the hyper-plane that separates data with maximum margin by categories. In our study, the sign of the sepsis severity score (Score) in Eq. 1 can predict the category for a patient. The magnitude of the Score represents the distance from the decision boundary and indicates the severity of sepsis. A large positive Score indicates critical severity.

Based on the fact that patients were hospitalized during the study, the longitudinal measurements should show a decrease in the number of patients in the PICU severe sepsis

group. Fig. 4 shows that Scores in the PICU severe sepsis group are notably separated from the PICU/sepsis group for the first two days after admission. After two days, the Scores in the PICU severe sepsis group decrease and collapse with those from the PICU/sepsis group indicating the effectiveness of the treatment.

## Comparison with the Random Forests learning method

Random forests (RF) [28] is an ensemble method [21, 22], which grows multiple classification and regression trees (CART) [29] for prediction. Every tree in the forests is constructed by a random selected bootstrap training set with replacement [21]. The splitting criteria for every decision node in a tree are also chosen from a random subset of the features without replacement. With the replacement from the original data, about two-third of the samples are used to construct a tree [21]. The out-of-bag (OOB) data, which are not chosen in the construction, are then used to estimate the prediction accuracy and the importance of the features [28, 30]. Unlike a linear SVM, which constructs a hyper-plane to classify the data, a tree is a hierarchical classification procedure, which recursively partitions the data to increase the purity of the nodes with respect to the outcome [29].

RF provides two measures, the mean decrease in accuracy (MDA) and mean decrease in the Gini index [28, 30], to estimate the importance of the features. In our study, the MDA is chosen to estimate the feature importance since the decrease in the Gini index is not as reliable as MDA [30, 31]. By randomly permuting the values of a given feature in the OOB data for each tree, RF measures the accuracy difference between untouched and permuted OOB data. The average of this accuracy difference over all trees in the forest is the MDA for the given feature. The MDA is the average increase in misclassification rate due to the permutations. The larger the MDA the more important the corresponding feature is with respect to the outcome.

Following Ref. [28], we construct a forest with 1,000 trees to estimate the MDA for biomarkers. We generated two RF: one for which Ang-2/Ang-1 is excluded (Fig. 5(A)) or included (Fig. 5(B)). Because of the interaction of Ang-2, Ang-1, and Ang-2/Ang-1, the existence of Ang-2/Ang-1 suppresses the importance of Ang-2 and Ang-1. However, both CCA and FS methods tend to select the combination of Ang-2 and Ang-1 as the most predictive feature. We notice that  $\text{HCO}_3$  is considered important for all three methods, which suggests  $\text{HCO}_3$  is also an important biomarker. The RF ranked biomarkers based on the importance are also shown in Table 1.

We also constructed a SVM ensemble using the RF-selected subset for each step  $k$  in Fig. 2 for comparison. Similar to the CCA-selected subset in Fig. 2, all prediction measures saturate at  $k = 4$  and decrease for  $k > 4$ . We find that the RF-selected optimal subset, Ang-2/Ang-1,  $\text{HCO}_3$ , Ang-2, Ang-1 at  $k = 4$ , have comparable prediction performance with the CCA-selected optimal subset at  $k = 3$ .

## Discussion

In this study, we employed machine learning approach to analyze the clinical data of children with severe sepsis using feature selection methods, such as CCA, SVM, FS and RF. Feature selection methods are helpful in identifying biomarkers with minimum redundancy that can be useful in clinical diagnosis. Our multivariate feature selection methods select the combination of Ang-1, Ang-2, and  $\text{HCO}_3$  as the optimal biomarkers for our data set. We demonstrated that this optimal combination of biomarkers significantly outperformed each single biomarker and all other combinations with redundant or irrelevant biomarkers for all statistical measures.

Our work [5, 7], and that of others, has shown the biological plausibility and clinical relevance of Ang-2 and Ang-1 levels in PICU patients with severe sepsis. It is interesting to note that combining Ang-2 and Ang-1 with a well-established (and routinely measured) indicator of an imbalance in the acid-base levels performs much better than other scoring systems that are more complex (for example, PIM-2 [13]).

Our data driven approach indicates that there is an optimal set of biomarkers for diagnosing severe sepsis. We have demonstrated that the use of additional biomarkers actually reduces the quality of the diagnostic scoring system. This is a potentially important observation in the sense that it suggests that more feature rich data may not be helpful, but actually harmful.

In addition, a sepsis severity score function (Eq. 1) using this optimal combination of biomarkers was constructed by the SVM ensemble method. With this function, we can interpret the relation between these three biomarkers and the sepsis severity from the associate weight,  $w_i$  [25]. Given that these relations have also been observed in clinical studies [3–5] we assert that our methodology is useful from the perspective of unbiased data analytical confirmation. It also holds promise for the discovery of novel biomarkers.

The proposed sepsis severity score for each sample is also evaluated during the treatment. The patients in PICU severe sepsis have significantly high severity scores after admission. After two days treatment, the severity scores for each patient decline and collapse to match patients without severe sepsis. Based on the fact that all patients survived hospitalization, the change in the longitudinal measurements of this score function validates the robustness and effectiveness of this function as regards its potential utility at different stages of treatment.

It has been observed that single biomarkers, in isolation, have limited diagnostic capacity [1]. This study supports this conclusion. Our analysis strongly supports the conclusion that a combination of different biomarkers is more effective, i.e., using multiple biomarkers for diagnosis is superior to drawing conclusions from single biomarkers. The rationale for this observation may be that the biomarkers are not independent of each other but, as we have shown with our canonical correlation analysis, are correlated in groups. The identification of an optimal combination of biomarkers allows clinician to focus on a small subset of indicators thus simplifying the diagnosis of sepsis in children with a spectrum



of severities.

In conclusion, based on three different models, we found that a linear additive combination of 3 biomarkers, namely Ang-2, Ang-1 and  $\text{HCO}_3$  provides a robust prediction of sepsis severity in patients admitted to the PICU. Additional independent studies are needed to confirm or refute the clinical utility of our biomarker combination for sepsis severity prediction. The collection of data sets with larger sample sizes would also be very useful for validating our statistical study.

## Methods

### Data Preprocessing

Our dataset (input), a  $n \times p$  real-valued matrix  $x$ , contains  $n = 45$  attributes and  $p = 19$  biomarkers. Since the range of values of the biomarkers varies widely, it should be normalized so that each biomarker contributes approximately proportionately. We normalized  $x$  to have zero mean and unit standard deviation for each biomarker [32]:

$$x_{\text{norm}} = \frac{x - \bar{x}}{\sigma(x)}, \quad (2)$$

where  $x_{\text{norm}}$  is a  $n \times p$  matrix,  $\bar{x}$  and  $\sigma(x)$  are the mean value and standard deviation of  $x$  for each biomarker. We also assigned each attribute  $i = 1, \dots, n$ , a sepsis severity score,  $y_i$ .  $y_i = 0$  is given to each in the PICU/sepsis group and  $y_i = 1$  for the PICU severe sepsis group.

### Canonical Correlation Analysis

CCA finds linear combinations of variables between two sets of data,  $x$  and  $y$  in our study, which have maximum correlation with each other [16, 17]. Here we selected the optimal subset of biomarkers  $x$  that has the maximum correlation with  $y$  for  $k = 1, \dots, p$ , by calculating the correlations between all possible  $k$ -combinations of  $x$  and  $y$ . The results are displayed in Table 1.

### Linear Support Vector Machines

In machine learning, a linear support vector machine (SVM) is a learning model used for classification and regression analysis [33]. A SVM model separates two categories by a hyper-plane that has maximum margin for a given training dataset. New attributes are predicted to belong to a category based on which side of the hyper-plane they fall on.

The search of this hyper-plane can be translated into the following optimization problem:

$$\begin{aligned}
 & \text{Minimize } \|w\|_1 + C_+ \sum_{i:y_i=+1} \xi_i + C_- \sum_{j:y_j=-1} \xi_j \\
 & \text{subject to} \\
 & \quad w^T x_i + b + \xi_i \geq 1, y_i = +1, \\
 & \quad w^T x_j + b - \xi_j \leq -1, y_j = -1, \text{ and} \\
 & \quad \xi \geq 0,
 \end{aligned} \tag{3}$$

where  $\|w\|_1 = \sum_i |w_i|$  is the 1-norm of a vector, which induces the sparsity in the weight vector  $w$  [24].

## Ensemble Method

Due to the limited size and noise of our data, we follow the training procedure in Ref. [21]. A random one-third of the data is selected as test set,  $T$ . The remaining data is used as training set,  $L$ . Bagging is used to construct the classifiers ensemble. Each new training set,  $L_i$ , is drawn, with replacement, from the original training set,  $L$ . Then a classifier, SVM or tree, is constructed on this new training set,  $L_i$ . In this study, we construct a classifiers ensemble 50 times,  $i \in \{1, \dots, 50\}$ . The final classification is obtained by calculating the mean of the ensemble of 50 classifiers. This procedure is repeated 100 times and statistical measures on  $T$  are averaged.

## Calculation of Statistical Measures

TPR, TNR, NPV, and PPV are statistical measures of the predictive performance of a binary classification test.

TPR (or sensitivity) measures the proportion of actual positives that are correctly identified.

TNR (or specificity) measures the proportion of actual negatives that are correctly identified.

PPV (or precision) measures the proportion of positives that are true positive.

NPV measures the proportion of negatives that are true negatives.

These statistical measures are calculated for each one of the 100 random divisions of test sets  $T$  by the classifier built on the bootstrap aggregation method. Their mean and standard error are calculated from the groups obtained from the 100 random divisions.

## Acknowledgments

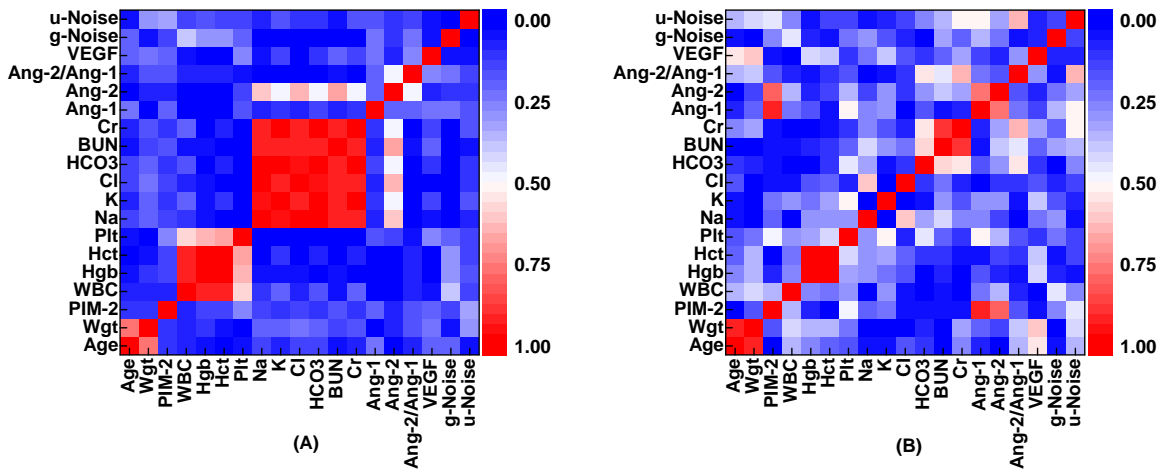
## References

1. Pierrakos C, Vincent JL, et al. (2010) Sepsis biomarkers: a review. *Crit Care* 14: R15.
2. Samraj RS, Zingarelli B, Wong HR (2013) Role of biomarkers in sepsis care. *Shock* 40: 358–365.
3. Ricciuto DR, dos Santos CC, Hawkes M, Toltl LJ, Conroy AL, et al. (2011) Angiopoietin-1 and angiopoietin-2 as clinically informative prognostic biomarkers of morbidity and mortality in severe sepsis. *Critical care medicine* 39: 702–710.
4. Fiusa M, Costa-Lima C, de Souza GR, Vigorito AC, Aranha F, et al. (2013) A high angiopoietin-2/angiopoietin-1 ratio is associated with a high risk of septic shock in patients with febrile neutropenia. *Crit Care* 17: R169.
5. Giuliano Jr JS, Tran K, Li FY, Shabanova V, Tala JA, et al. (2014) The temporal kinetics of circulating angiopoietin levels in children with sepsis. *Pediatric Critical Care Medicine* 15: e1–e8.
6. Siner JM, Bhandari V, Engle KM, Elias JA, Siegel MD (2009) Elevated serum angiopoietin 2 levels are associated with increased mortality in sepsis. *Shock* 31: 348–353.
7. Giuliano Jr JS, Lahni PM, Harmon K, Wong HR, Doughty LA, et al. (2007) Admission angiopoietin levels in children with septic shock. *Shock* (Augusta, Ga) 28: 650.
8. Parikh SM (2013) Dysregulation of the angiopoietin–tie-2 axis in sepsis and ards. *Virulence* 4.
9. El-Solh AA, Jaoude PA, Porhomayon J (2010) Bicarbonate therapy in the treatment of septic shock: a second look. *Internal and emergency medicine* 5: 341–347.
10. Wong HR, Salisbury S, Xiao Q, Cvijanovich NZ, Hall M, et al. (2012) The pediatric sepsis biomarker risk model. *Crit Care* 16: R174.
11. Wong HR, Weiss SL, Giuliano Jr JS, Wainwright MS, Cvijanovich NZ, et al. (2014) Testing the prognostic accuracy of the updated pediatric sepsis biomarker risk model. *PLOS ONE* 9: e86242.

12. Wong HR, Weiss SL, Giuliano Jr JS, Wainwright MS, Cvijanovich NZ, et al. (2014) The temporal version of the pediatric sepsis biomarker risk model. *PloS one* 9: e92121.
13. Slater A, Shann F, Pearson G (2003) Pim2: a revised version of the paediatric index of mortality. *Intensive care medicine* 29: 278–285.
14. Goldstein B, Giroir B, Randolph A, et al. (2005) International pediatric sepsis consensus conference: Definitions for sepsis and organ dysfunction in pediatrics. *Pediatric critical care medicine* 6: 2–8.
15. Wang K, Bhandari V, Chepustanova S, Huber G, Stephen O, et al. (2013) Which biomarkers reveal neonatal sepsis? *PloS one* 8: e82700.
16. Mardia KV, Kent JT, Bibby JM (1980) *Multivariate analysis* .
17. Tofallis C (1999) Model building with multiple dependent variables and constraints. *Journal of the Royal Statistical Society: Series D (The Statistician)* 48: 371–378.
18. Björck A, Golub GH (1973) Numerical methods for computing angles between linear subspaces. *Mathematics of computation* 27: 579–594.
19. Efroymson M (1960) Multiple regression analysis. *Mathematical methods for digital computers* 1: 191–203.
20. Sjöstrand K, Clemmensen LH, Larsen R, Ersbøll B (2012) Spasm: A matlab toolbox for sparse statistical modeling. *Journal of Statistical Software* Accepted for publication .
21. Breiman L (1996) Bagging predictors. *Machine learning* 24: 123–140.
22. Dietterich TG (2000) Ensemble methods in machine learning. In: *Multiple classifier systems*, Springer. pp. 1–15.
23. Kim HC, Pang S, Je HM, Kim D, Yang Bang S (2003) Constructing support vector machine ensemble. *Pattern recognition* 36: 2757–2767.
24. Mangasarian OL (1999) Arbitrary-norm separating plane. *Operations Research Letters* 24: 15–23.
25. Mladenić D, Brank J, Grobelnik M, Milic-Frayling N (2004) Feature selection using linear classifier weights: interaction with classification models. In: *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, pp. 234–241.

26. van der Heijden M, Pickkers P, van Nieuw Amerongen GP, van Hinsbergh VW, Bouw MP, et al. (2009) Circulating angiotensin-2 levels in the course of septic shock: relation with fluid balance, pulmonary dysfunction and mortality. *Intensive care medicine* 35: 1567–1574.
27. David S, Mukherjee A, Ghosh CC, Yano M, Khankin EV, et al. (2012) Angiotensin-2 may contribute to multi-organ dysfunction and death in sepsis. *Critical care medicine* 40: 3034.
28. Breiman L (2001) Random forests. *Machine learning* 45: 5–32.
29. Breiman L, Friedman J, Stone CJ, Olshen RA (1984) *Classification and regression trees*. CRC press.
30. Pang H, Lin A, Holford M, Enerson BE, Lu B, et al. (2006) Pathway analysis using random forests classification and regression. *Bioinformatics* 22: 2028–2036.
31. Breiman L (2002) *Manual on setting up, using, and understanding random forests v3. 1*. Retrieved October 23: 2010.
32. Morik K, Brockhausen P, Joachims T (1999) Combining statistical learning with a knowledge-based approach-a case study in intensive care monitoring. In: *Machine Learning-International Workshop Then Conference*. Morgan Kaufmann Publishers, Inc., pp. 268–277.
33. Vapnik V (2000) *The nature of statistical learning theory*. springer.

## Figure Legends



**Figure 1. Heatmaps of pairwise correlations.** Heatmaps of all pairwise correlations between the 17 variables (plus two noise samples) for patients in the (A) PICU/sepsis and (B) PICU severe sepsis groups. The color scale from blue to red indicates increasing correlations between the pair of biomarkers at the corresponding locations on the horizontal and vertical axes.

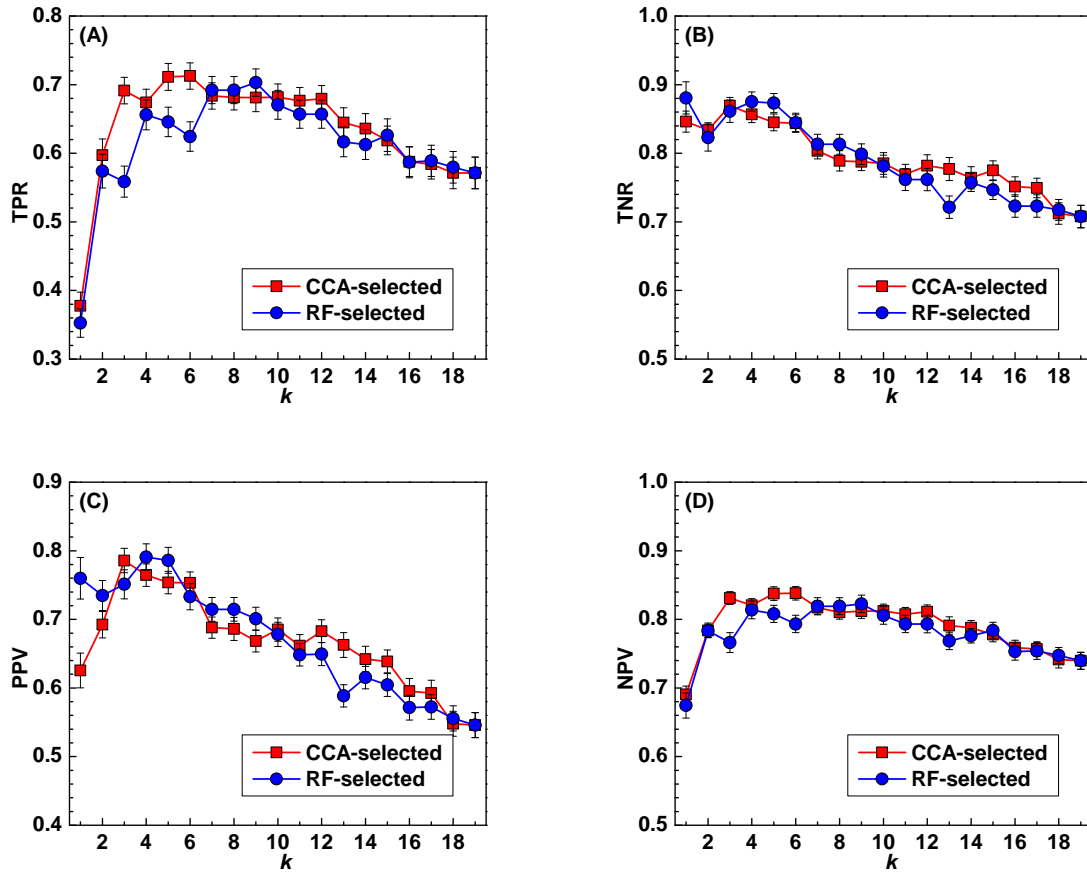
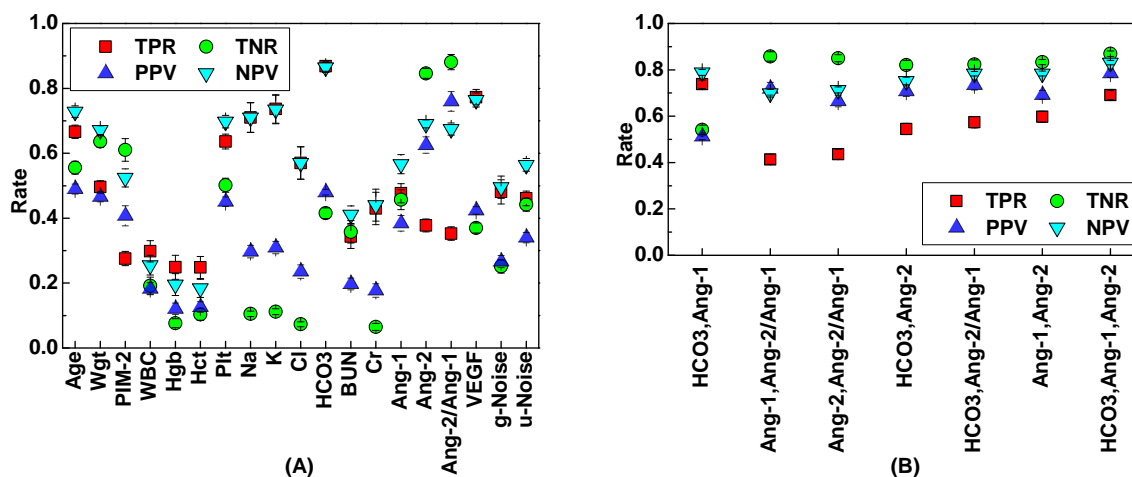
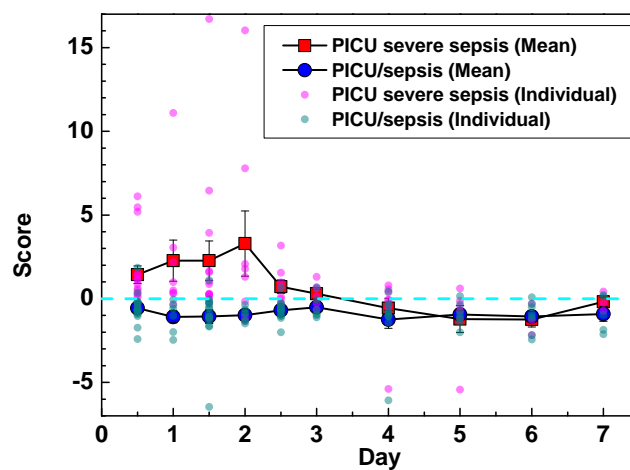


Figure 2. Prediction measures obtained from the Support Vector Machine (SVM) using the  $k$ -combinations selected by Canonical Correlation Analysis (CCA) and Random Forests (RF) methods. The prediction measures (A) true positive rate (TPR), (B) true negative rate (TNR), (C) positive predictive value (PPV), and (D) negative predictive value (NPV) are shown for each step  $k$ . For each  $k$ , a SVM ensemble with bagging is constructed based on the CCA- and RF-selected subset of biomarkers.

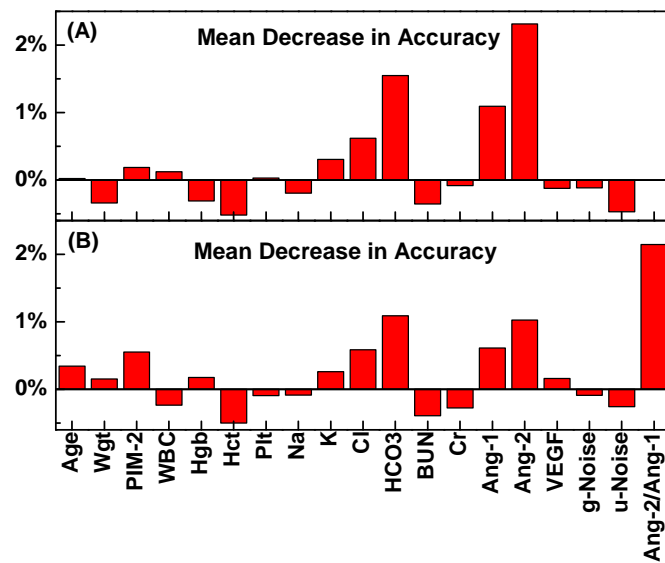


**Figure 3. Prediction measures for single and pairs of biomarkers from the Support Vector Machine (SVM).** True positive rate (TPR), true negative rate (TNR), positive predictive value (PPV), and negative predictive value (NPV) are shown for (A) each single biomarker and (B) all pairwise combinations of Ang-1, Ang-2, HCO<sub>3</sub> and Ang-2/Ang-1. The prediction measures for the CCA-selected optimal subset of biomarkers at  $k = 3$  (Ang-2, Ang-1, and HCO<sub>3</sub> are also shown in (B) for comparison.



**Figure 4. Longitudinal measurements of the sepsis severity score.** The sepsis severity scores (Score) for patients from the PICU/sepsis group and the PICU severe sepsis during the 7 days of illness. Both the mean and individual severity scores are plotted.





**Figure 5.** Measures of the biomarker importance obtained from the **Random Forests method**. Mean Decrease in Accuracy (MDA) are shown for biomarkers in (A) without Ang-2/Ang-1 and (B) with Ang-2/Ang-1 using the Random Forests method with 1,000 trees for each.

## Tables

**Table 1. Stepwise Biomarker Selection using Canonical Correlation Analysis, Forward Selection and Random Forests.**

Dim	Corr	Entering	Leave	Forward Selection	Random Forests
1	0.3811	Ang-2		Ang-2	Ang-2/Ang-1
2	0.4772	Ang-1		Ang-1	HCO3
3	0.5501	HCO3		HCO3	Ang-2
4	0.5842	Plt		Plt	Ang-1
5	0.6079	Age		Age	Cl
6	0.6183	Cl		WBC	PIM-2
7	0.6221	BUN, Hct, WBC	Cl, HCO3	Hct	Age
8	0.6286	VEGF		BUN	K
9	0.6311	PIM-2		VEGF	Hgb
10	0.6359	Cl, HCO3	PIM-2	PIM-2	VEGF
11	0.6395	Cr, Wgt	Age	g-Noise	Wgt
12	0.6409	Hgb, Na, Age	Cl, Wgt	Cl	Na
13	0.6414	Ang-2/Ang-1		Cr	g-Noise
14	0.6419	Wgt		u-Noise	Plt
15	0.6424	PIM-2		Ang-2/Ang-1	WBC
16	0.6427	Cl, u-Noise	Na	Hgb	u-Noise
17	0.6429	K		Wgt	Cr
18	0.6429	Na, g-Noise	K	K	BUN
19	0.6430	K		Na	Hct

We apply Canonical Correlation Analysis for all possible  $k$ -combinations ( $k = 1, \dots, 19$ ) to determine the subset of  $k$  biomarkers with the highest correlation with the sepsis severity score. The ‘Enter’ column indicates the biomarker that is added to achieve the highest correlation at each  $k$ . The ‘Leave’ column indicates the biomarker that is eliminated from the combination at each  $k$ . A biomarker will stay in the combination until it occurs in ‘Leave’ column. The ‘Forward Selection’ column gives the biomarker selected by the Forward Selection method when applied one biomarker at a time. The ‘Random Forests’ column gives the biomarker ranked by the mean decrease in accuracy measured by the Random Forests method.

**Table 2. Prediction measures for single variable from Support Vector Machine.**

Variable	TPR	TNR	PPV	NPV
Age	0.666	0.555	0.490	0.728
Wgt	0.496	0.636	0.466	0.671
PIM-2	0.276	0.611	0.407	0.524
WBC	0.298	0.192	0.183	0.255
Hgb	0.249	0.076	0.121	0.195
Hct	0.248	0.104	0.126	0.184
Plt	0.636	0.501	0.451	0.697
Na	0.710	0.105	0.297	0.710
K	0.737	0.112	0.309	0.735
Cl	0.570	0.073	0.236	0.570
HCO <sub>3</sub>	0.868	0.415	0.480	0.865
BUN	0.343	0.358	0.197	0.411
Cr	0.430	0.065	0.177	0.440
Ang-1	0.477	0.457	0.384	0.566
Ang-2	0.378	0.846	0.625	0.690
Ang-2/Ang-1	0.353	0.881	0.760	0.675
VEGF	0.773	0.370	0.424	0.764
g-Noise	0.481	0.251	0.266	0.496
u-Noise	0.461	0.442	0.340	0.564

True positive rate (TPR), true negative rate (TNR), positive predictive value (PPV), and negative predictive value (NPV) are shown for each single variable.

**Table 3. Parameters for the decision function that includes the CCA-selected optimal subset of biomarkers at  $k = 3$ .**

i	Biomarker	Mean $\bar{x}_i$	Standard Deviation $\sigma_i$	Weight $w_i$ ( $b = 0.313$ )	Standard Error of Weight $e_i$
1	Ang-2	8518.1	13264	1.994	0.065
2	Ang-1	2649.2	4008.9	-1.396	0.050
3	HCO <sub>3</sub>	27.270	24.361	-1.340	0.072

The values of the weights  $w_i$ , errors  $e_i$ , means  $\bar{x}_i$ , and standard deviations  $\sigma_i$  for the biomarkers in Eq. (1).