Article

K-Means Clustering for Shock Classification in Pediatric Intensive Care Units

María Rollán-Martínez-Herrera [1,2,3,*], Jon Kerexeta-Sarriegi [3,4,5], Javier Gil-Antón [1,2], Javier Pilar-Orive [1,2] and Iván Macía-Oliver [3,4]

1 Cruces University Hospital, 48903 Barakaldo, Spain
2 Biocruces Bizkaia Health Research Institute, 48903 Barakaldo, Spain
3 Vicomtech Foundation, Basque Research and Technology Alliance (BRTA), Mikeletegi 57, 20009 Donostia, Spain
4 Biodonostia Health Research Institute, 20018 Donostia, Spain
5 Computational Intelligence Group, Computer Science Faculty, University of the Basque Country, UPV/EHU, 48940 Donostia, Spain
* Correspondence: meroll16@gmail.com

Abstract: Shock is described as an inadequate oxygen supply to the tissues and can be classified in multiple ways. In clinical practice still, old methods are used to discriminate these shock types. This article proposes the application of unsupervised classification methods for the stratification of these patients in order to treat them more appropriately. With a cohort of 90 patients admitted in pediatric intensive care units (PICU), the k-means algorithm was applied in the first 24 h data since admission (physiological and analytical variables and the need for devices), obtaining three main groups. Significant differences were found in variables used (e.g., mean diastolic arterial pressure $p < 0.001$, age $p < 0.001$) and not used for training (e.g., EtCO2 min $p < 0.001$, Troponin max $p < 0.01$), discharge diagnosis ($p < 0.001$) and outcomes ($p < 0.05$). Clustering classification equaled classical classification in its association with LOS ($p = 0.01$) and surpassed it in its association with mortality ($p < 0.04$ vs. $p = 0.16$). We have been able to classify shocked pediatric patients with higher outcome correlation than the clinical traditional method. These results support the utility of unsupervised learning algorithms for patient classification in PICU.

Keywords: shock; pediatric; unsupervised learning; k-means; stratification

1. Introduction

Shock is described as an inadequate oxygen supply to the tissues and can be classified in multiple ways. There are four classic pathophysiological variants: cardiogenic, hypovolemic, distributive, and obstructive, which explain the hemodynamic situation of the patient and are therefore useful for establishing therapy. Shock can also be classified according to its etiology: septic, anaphylactic, hemorrhagic, etc. The problem with these classifications is they are often confusing, since each pathophysiological type has different etiological varieties and each of the etiological varieties may correspond to different pathophysiological types [1–5].

When faced with a new case of shock, it is necessary to simultaneously evaluate multiple physiological variables to determine the pathophysiological variety, in order to establish a targeted therapy [3]. However, despite the time that has been spent researching hemodynamic monitoring, there are still no adequate techniques to make an accurate shock classification [4–6], which negatively affects the clinical management.

Multiple studies have been performed, and classical statistical analysis has been used to determine the type of shock in each patient. However, no single solution to this historical problem has yet been found. Currently, there are multiple models of biological data analysis using artificial intelligence techniques based on machine learning [7]. The use of supervised

learning for patient categorization would imply the assumption that the ground truth of the disease classification is already perfect, but in practical terms, this classification may not be optimal or be outdated. Therefore, in clinical practice, similarities between patients are continually encountered, which in the long run sometimes end up defining new categories.

The present study proposes that the use of unsupervised learning techniques based on the multiple variables routinely collected in pediatric intensive care units (PICU) could help to find patterns that clinicians may have overlooked, and that may define new classifications of types of shock in the pediatric setting. The use of clustering algorithms for the classification of shock could not only be applied globally but could also be used locally. It could be useful to study the characteristics of each hospital and classify patients into clinical groups to study treatment and evolution, similarly to that performed at the level of microbiological flora by analyzing local behavior and resistances.

2. Materials and Methods

A retrospective observational study was carried out for the development of a computational model for the classification of types of shock in the PICU of Cruces University Hospital. Clinical and analytical data were collected for all pediatric patients (0 to 14 years) diagnosed with any type of shock since the implementation of "IntelliSpace Critical Care and Anesthesia", in 2012.

Data were collected in a computerized manner through the program's database: 180 patients had been diagnosed with shock, but only 100 had data. Hourly data of physiological, gasometric, and analytical variables were obtained. In addition, devices, age, weight, length of stay, diagnostic at discharge, and discharge reason, were recorded.

From all the information available at the beginning, a first filtering was made by removing those variables with more than 80% of missing values and those with no clinical relevance (such as the number of central venous catheter lumens, and urinary catheter size, etc.). Then, patients with missing values in the heart rate, respiratory rate, and pressure columns were eliminated. Afterward, the resulting dataset comprised 90 patients. In order to achieve higher clinical value in the classification, only data from the first 24 h of admission were selected, and data preceding death in less than 48 h were eliminated.

For those variables that depend on age, the values were adjusted with the z-score for that age (weight [8,9], blood pressure (using the p. 50 for height) [10], and heart and respiratory rate [11]).

After that, the mean, minimum (min), and maximum (max) values of each monitored numerical variable were calculated. A classification of the patients was performed by clustering.

The following statistical techniques were used. Lilliefors test was used to analyze normality. Normal variables were described by mean, non-normal variables by the median, and qualitative variables by proportion. Dispersion measures have been estimated by the confidence interval of 95%. For quantitative variables with 2 groups, the *t*-test (with Welch correction if homoscedasticity was not met) or Wilcoxon test were used depending on normality. In quantitative variables with more than 2 groups, ANOVA or Kruskal–Wallis were used depending on normality, and post hoc analyses were performed with Tukey and the Mann–Whitney test with significance correction, respectively. Chi-square was used for qualitative variables. In the case of length of stay, Log-Rank was also used, with post hoc analysis using Bonferroni correction. A $p < 0.05$ was considered the cut-off point for statistical significance.

K-means was the algorithm selected for clustering [12]. This algorithm classifies individuals into multiple groups so that individuals within the same group are as similar as possible, while individuals from different groups are as different as possible. This similarity is estimated using the Euclidean distance [13], so that if two patients have very different values for a variable, the distance will be high, and vice versa. Each group is represented by its centroid, which corresponds to the mean of all individuals in that group, the initial centroids are randomly selected by the algorithm, and this may influence the results, which is why

25 initial configurations were attempted. As a characteristic of the k-means algorithm, it is necessary to define both the grouping variables and the number of clusters.

For the clustering classification, only the relevant physiological variables with less than 45% of missing values were used. To study the optimal number of clusters, the elbow method and the average silhouette method were used, and several tests were performed to check the spatial distribution of the groups and the number of patients per group. Three was taken as the optimal number of clusters.

Once the grouping variables and the number of clusters were selected, the data were prepared: each missing value was estimated by averaging the remaining individuals, and then the data were standardized with z-score [14] to make all variables comparable.

Although it has been shown that good results can be achieved by giving more weight to some variables than others [15], in this study, all variables have been treated with equal importance to avoid both clinician subjectivity and overfitting when assigning weights (due to the size of the database).

Once the clusters were selected, the clinical significance of those clusters was sought as follows:

- The characteristics of each group were studied to determine whether there were differences between them.
- The correlation between the unsupervised classification and the discharge diagnosis was studied.
- It was assessed whether the classification was related to the outcomes (mortality and length of stay).
- It was tested whether the new classification had a greater association with outcomes than the classic classification.

All analyses were performed with R-Studio; tables and graphs were made with R-studio and Microsoft Excel.

3. Results

K-means classification was performed using the variables indicated in Table A1 (shown in the appendix due to its size). The following clusters were obtained: 46 patients, 18 patients, and 26 patients (Figure 1).
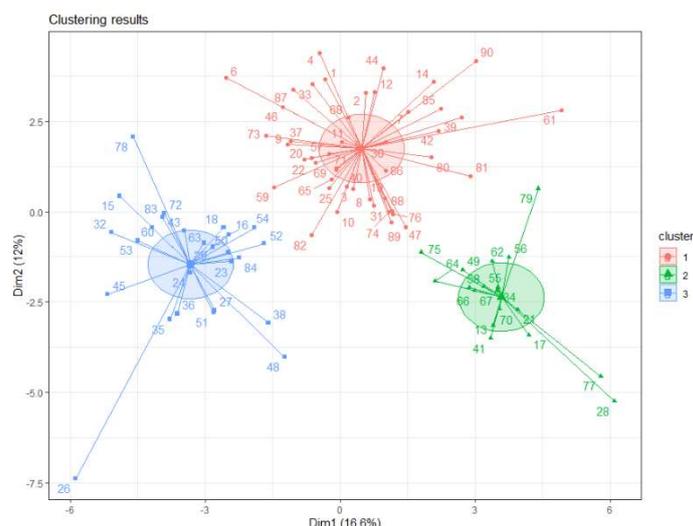
Figure 1. Two—dimensional representation of the clusters based on principal component analyses.

3.1. Analysis of Variables Used for Clustering

Each of the variables was compared between the different clusters (Table A1). Among all the variables, the need for ECMO was the one with the greatest weight, so the clustering was performed again without it. The clusters obtained were 45, 19 and 26 patients, almost the same distribution as before (ECMO: 0%, 79% and 2%, respectively, $p < 0.001$).

Cluster 1 was the cluster with the highest proportion of female patients (56%), the one with the lowest weight for age, and the one with the lowest mean arterial pressure and diastolic pressure. The patients in this cluster had the lowest carboxyhemoglobin, the highest calcium ion, the lowest creatinine, and the highest number of neutrophils. The cluster with the highest proportion of intracranial catheters and thermal blankets and the lowest proportion of patients with hemodiafiltration.

Cluster 2 was characterized as the one with lower mean age, heart and respiratory rate, systolic blood pressure, oxygen saturation, venous oxygen saturation, temperature, and higher inspired oxygen fraction. It was the group of patients with higher daily diuresis, wider capillary glycemia range, lower calcium ion levels, higher phosphate levels, lower C-reactive protein levels, higher lymphocytosis and lower neutrophilia. It was the cluster with the highest proportion of patients with ECMO, mechanical ventilation and hemodiafiltration.

Cluster 3 was the one with older patients and higher weight for age. These patients had the highest heart rate and respiratory rate, highest blood pressure, lowest daily diuresis, highest oxygen saturation and lowest inspired oxygen fraction, highest temperature, venous oxygen saturation and carboxyhemoglobin. This was the group with higher capillary glycemia, creatinine, lower phosphate levels, higher C-reactive protein, neutropenia and lymphopenia. In addition, this group had the lowest proportion of patients on mechanical ventilation, ECMO and heat blanket.

Clusters 1 and 3 were the most similar; however, cluster 3 presented higher age ($p < 0.001$), higher weight for age ($p < 0.006$) and higher proportion of males ($p < 0.04$). It presented higher levels in the average, maximum and minimum values of all tensions ($p < 0.001$). Finally, cluster 3 presented lower lymphocytosis and neutrophilia ($p < 0.001$) and higher lymphopenia and neutropenia ($p < 0.001$), in addition to a higher C-reactive protein ($p < 0.001$).

3.2. Analysis of Variables Not Used for Clustering

Each variable was compared among the three clusters (Table A1). Exhaled CO_2 pressure showed important differences, with maximum levels in cluster 1, followed by cluster 3; cluster 2 presented much lower levels than the other two. Troponin also presented significant differences, with the highest level in cluster 2, followed by cluster 1 and with minimum levels in cluster 3.

### 3.3. Relationship between Clustering and the Classic Classification

Cardiogenic shock presented different proportions between groups 1 and 2 ($p < 0.001$) and 2 and 3 ($p < 0.001$); however, groups 1 and 3 did not present differences. Inflammatory shock presented significant differences between all combinations (1 vs. 2, $p < 0.03$; 2 vs. 3, $p < 0.001$; 3 vs. 1, $p < 0.04$). Hypovolemic shock was not specifically associated with any group. Cluster 2 had the highest number of postoperative cardiac surgery patients (44%, $p < 0.001$), while cluster 3 had the highest number of oncologic patients (31%, $p = 0.003$). In Figure 2 the absolute number of patients per classic classification of each cluster is shown.
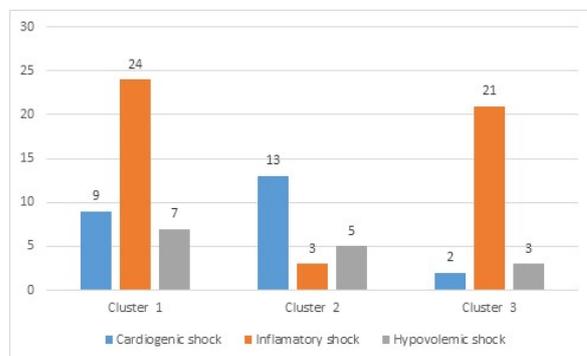
Figure 2. Number of patients with each type of shock in every cluster.

3.4. Analysis of Outcomes According to Clustering

Both lengths of stay and death were chosen as outcomes. In the case of length of stay (Figure 3, part A), significant differences were found between the three groups ($p < 0.03$), mainly because of the differences between groups 2 and 3 ($p < 0.02$). Groups 1 and 2 did not present significant differences ($p < 0.11$), and neither did groups 1 and 3 ($p < 0.28$). Using Log-Rank, differences were also found between the three groups ($p = 0.01$), mainly due to groups 2 and 3 ($p < 0.02$); between groups 1 and 2, no differences were found ($p = 0.18$), nor between groups 1 and 3 ($p = 0.8$). Significant differences were observed between clusters in survival ($p < 0.04$), also mainly due to differences with cluster 2. Between clusters 1 and 2, there were significant differences ($p < 0.02$), and the same was true between clusters 2 and 3 ($p < 0.03$). However, between groups 1 and 3, there were no differences ($p = 1$).
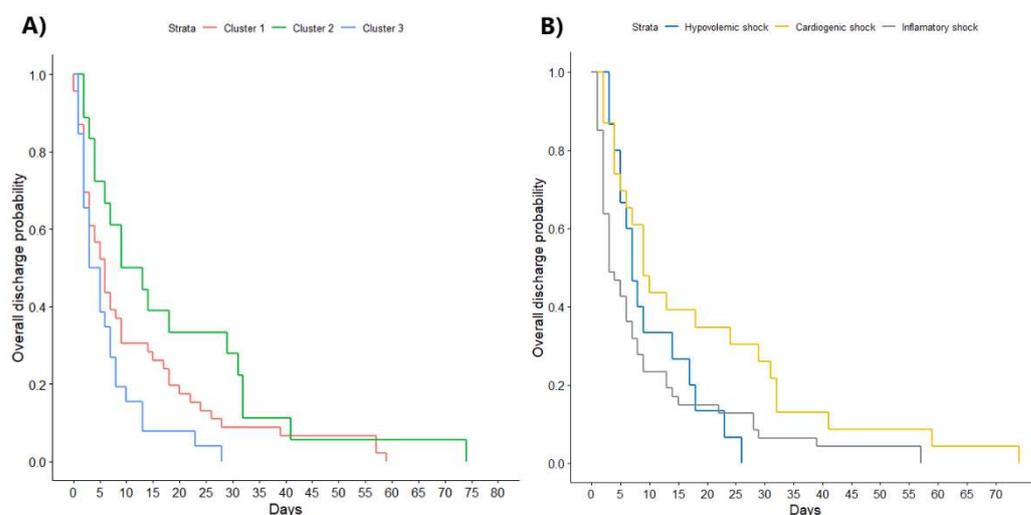
Figure 3. Kaplan–Meier curves for the length of stay in the intensive care unit (**A**) according to clusters and (**B**) according to classic classification.

3.5. Prediction of Outcomes by Classic Classification

Patients with cardiogenic shock had significantly longer length of stay than patients without cardiogenic shock (Wilcoxon *p*-value < 0.02, Log-Rank *p*-value = 0.01). Patients with inflammatory shock had shorter length of stay than patients without it (Wilcoxon *p*-value < 0.02, Log-Rank *p*-value = 0.05). As for patients with hypovolemic shock, there was no significant association with length of stay (Figure 3, part B).

Cardiogenic and inflammatory shock presented an inverse association between them, only one patient presented both types of shock, and none presented either of them ($p < 0.001$). Given this circumstance and the fact that it makes no sense for either type of shock to shorten the length of stay, it is most likely that the true determinant of length of stay is cardiogenic shock.

As for the probability of death, none of the three classic types of shock was associated with the survival rate (Table 1).

Cardiac postoperative patients had significantly longer length of stay than non-postoperative patients (9 vs. 5 days, Wilcoxon *p*-value < 0.03, Log-Rank *p*-value = 0.01), while oncologic patients did not show significant differences in length of stay with respect to non-oncologic patients (3 vs. 6 days, Wilcoxon *p*-value = 0.15, Log-Rank *p*-value = 0.1). Regarding mortality, no significant differences were found between groups in any case (cardiac surgery: 8% vs. non-cardiac surgery: 12%, $p = 1$; oncologic: 0% vs. non-oncologic: 13%, $p = 0.45$).

Globally, no significant differences were found in the LOS between patients that died and patients that survived (4.5 vs. 6 days, Wilcoxon *p*-value = 0.91, Log-Rank *p*-value = 0.8). However, analyzing cardiogenic shock patients, it was found that patients that died had a shorter LOS than survivors (4 vs. 13 days, Wilcoxon *p*-value = 0.04, Log-Rank

p-value = 0.004). Additionally, cluster 2 showed this behavior (4 vs. 21.5 days, Wilcoxon *p*-value = 0.02, Log-Rank *p*-value = 0.006). The other two types of shock and clusters 1 and 3 presented the opposite trend but with no significant differences.

Table 1. Association between each type of shock and both length of stay and death.

Types of Shock	No Cardiogenic	Cardiogenic	Wilcoxon/χ2 (*p* Value)	Log-Rank (*p* Value)
Median length of stay (days)	5 (3; 7)	9 (4; 24)	0.02	0.01
Exitus	0.08 (0.01; 0.14)	0.21 (0.03; 0.38)	0.16	-
	No inflammatory	**Inflammatory**	**Wilcoxon/χ2 (*p* value)**	**Log-Rank (*p* value)**
Median length of stay (days)	7.5 (5; 13)	3 (2; 6)	0.01	0.05
Exitus	0.14 (0.03; 0.25)	0.08 (0; 0.16)	0.58	-
	No hypovolemic	**Hypovolemic**	**Wilcoxon/χ2 (*p* value)**	**Log-Rank (*p* value)**
Median length of stay (days)	5 (3; 7)	7 (5; 14)	0.17	1
Exitus	0.11 (0.04; 0.18)	0.13 (0; 0.33)	1	-

4. Discussion

It is important to consider that the clustering classification has been performed with data of a specific hospital so that it is not generalizable. Post-operative care for congenital heart disease, the absence of cardiac transplantation or the presence of a pediatric oncology unit are some of the particularities that determine the type of patients admitted with shock in this hospital, and therefore the classification made. Therefore, the proposal in this article is not the generalization of the model presented, but the demonstration that clustering algorithms (specifically k-means) can be useful for classifying shocked pediatric patients, and that they can even be more accurate than clinical classification.

Data available were mainly from two groups according to the classical classification, patients with cardiogenic shock and patients with inflammatory shock. However, clustering classification was made in three groups, and it seems optimal in its two-dimensional representation (Figures 1 and 2).

The analysis of the different classificatory variables showed that some were more relevant than others. Specifically, ECMO was the variable with the greatest discriminatory weight, and therefore, a new analysis was performed after eliminating this variable. When this variable was eliminated, the groups barely changed, which ruled out the possibility that the good clustering results were due to this variable, and demonstrated the classificatory quality of the algorithm that could discriminate ECMO patients.

Previous studies have shown that clustering algorithms tend to group intensive care patients into physiologically similar groups [16]. Cluster 1 appeared to belong to a group of low-weight, hypotensive infants and young children, mainly with septic shock and some cardiogenic shock. Cluster 2 seemed to correspond to the most severe patients, infants with cardiogenic shock, in a large proportion of postoperative cardiac patients (44%, *p* < 0.001). Cluster 3 seemed to correspond to the least severe patients, neutropenic children, with apparent distributive septic shock, in a large proportion of oncologic patients (39%, *p* < 0.004).

The two main classical types of shock were associated with different clusters; cluster 2 had the highest proportion of cardiogenic shock (72%, *p* < 0.001) and inflammatory shock was mainly associated with clusters 1 (52%) and 3 (81%), but with significant differences between them (*p* < 0.04) and with differences between them and cluster 2 (*p* < 0.03). Considering that the clustering classification was performed only using data from the first 24 h, and the classical classification was performed based on the discharge diagnosis, these results are very promising, and they could suggest that clustering is able to discriminate patients earlier than the clinical eye.

As for the assessment of whether this new classification was able to predict outcomes, the length of stay was different among the three clusters (11 vs. 4 days, Log-Rank $p < 0.02$). Graphically (Figure 3), the Kaplan–Meier curves were different for each cluster of patients; therefore, the absence of differences between groups 1–2 and 2–3 could be due to the small amount of data. The prediction of mortality was also satisfactory, since groups 1 and 2 (7% vs. 33%, $p < 0.02$) and 2 and 3 (33% vs. 4%, $p < 0.03$) presented significant differences; groups 1 and 3, both having such a small proportion of deceased patients, did not present differences. These results are consistent with those of previous analyses, which also suggest that clustering can provide prognostic information [17].

The final objective of this study was to determine whether the use of clustering algorithms could be better than the use of clinical classifications. To do so, it was necessary to compare the outcome prediction ability of the two types of classifications, but is important to bear in mind that the clustering classification was performed with data from the first 24 h; moreover, clustering classification has fewer patients per group than the classical classification, so differences between groups were less likely to be significant.

Cardiogenic shock was associated with a longer length of stay (9 vs. 5 days, Log-Rank $p = 0.01$) with the same significance as cluster classification (6 vs. 11 vs. 4 days, Log-Rank $p = 0.01$); therefore, in this scenario, clustering did not improve the prediction of length of stay. The association with mortality by classical categorization was not significant in any case (cardiogenic shock 21% vs. non-cardiogenic shock 8%, $p = 0.16$); however, the association with mortality using clustering classification was significant (7% vs. 33% vs. 4%, $p = 0.003$); therefore, classification by clustering improved the prediction of mortality.

Although the field of artificial intelligence applied to medicine has been booming in recent years, most of the work focuses on the technical side of the algorithms. The results of the present study represent an approximation between the technical world and the clinical world. It seems that, even with such a small sample, the presented method is able to classify patients with a higher outcome correlation than the clinical method. Other studies have also presented positive results in the application of this type of clustering in other pathologies [17].

The capacity for early classification of shocked patients presents multiple clinical utilities. From the evolutionary point of view, a correct classification allows the clinician to identify the type of patient he/she is dealing with and, based on the historical behavior of similar patients, allows him/her to predict the evolution of the patient. From the therapeutic point of view, patients who behave the same physiologically tend to respond similarly to the same therapies, so a correct classification allows for improved treatment.

The classification obtained in the present study is not generalizable, but what is generalizable is the method applied to obtain it. However, it is necessary to consider some limitations. When data are collected digitally, sometimes they present incorrect recordings, which has to be cautiously interpreted [17–19], but this limitation is also present in clinical classification. Moreover, k-means algorithm requires pre-specifying a number of clusters, which means it is not absolutely automatic, it also needs the imputation of missing values [17] and it is very sensitive to outliers. Additionally, the size of the trained dataset is small and from a single hospital, so the results, pending validation, are not generalizable to other hospitals. Finally, the variables that have been used are static, i.e., they have been atemporalized, and thus, it would be interesting, as future work, to carry out this experiment but considering all the sequence of the data.

Despite these issues, the present results are still encouraging, and they support the utility of these methods in the classification of patients with shock in pediatric intensive care, something that could help to implement guided therapies [17]. If each hospital develops its own classification, which will be different according to their characteristics, or if a multicenter classification is performed, this may help to improve the clinical management of pediatric patients with shock.

5. Conclusions

In conclusion, the present study demonstrates the capacity of the k-means algorithm to correctly classify pediatric patients with shock and shows a promising future for the use of unsupervised machine learning techniques in pediatric critical care, something that could lead to an improvement in clinical management. However, further studies are needed to validate this method on a larger scale.

Appendix A

Table A1. Comparison of variables. Mean (mn) or median (md) of the 24 h first data (min, mean, max) per group according to the normality of quantitative data or proportion (p) of qualitative data, *p*-value of differences.

Variables used for clustering	Cluster 1 (Mean/Median (CI95%))	Cluster 2 (Mean/Median (CI95%))	Cluster 3 (Mean/Median (CI95%))	*p* Value
Female sex	p: 0.57 (0.42; 0.71)	p: 0.33 (0.09; 0.57)	p: 0.31 (0.12; 0.5)	0.06
Age in months	md: 33.63 (14.97; 55.17)	md: 3.25 (0.2; 8.77)	md: 137.13 (102.17; 169.43)	**<0.001**
Weight z-score for age	md: −1.3 (−2.18; −0.51)	md: −0.93 (−2.07; 0.16)	md: 0.03 (−0.82; 2.05)	**0.02**
Hearth rate z-score for age (first 24 h min)	md: −0.32 (−0.62; 0.35)	md: −2.59 (−5.97; −0.14)	md: −0.27 (−1.26; 0.42)	**0.01**
"(first 24 h mean)	mn: 1.82 (1.4; 2.24)	mn: 0.77 (−0.11; 1.64)	mn: 1.99 (1.41; 2.58)	0.93
"(first 24 h max)	md: 3.82 (3.44; 4.31)	md: 2.7 (1.85; 5.36)	md: 5.13 (4.39; 5.45)	**0.002**
Respiratory rate z-score for age (first 24 h min)	md: −2.33 (−2.78; −1.99)	md: −3.61 (−5.14; −2.67)	md: −1.06 (−2.33; −0.63)	**<0.001**
"(first 24 h mean)	mn: 1.07 (0.2; 1.95)	mn: −1.04 (−2.6; 0.51)	mn: 2.18 (0.87; 3.49)	0.32
"(first 24 h max)	mn: 6 (3.49; 8.51)	mn: 2.63 (−0.38; 5.63)	mn: 7.49 (5.01; 9.98)	0.59
Diastolic arterial pressure z-score for age (first 24 h min)	md: −0.15 (−0.15; 0.02)	md: −0.09	md: 0.05 (−0.1; 0.1)	**0.01**
"(first 24 h mean)	mn: 0.24 (0.19; 0.29)	mn: 0.44 (0.09; 0.78)	mn: 0.54 (0.48; 0.61)	**<0.001**
"(first 24 h max)	mn: 0.78 (0.7; 0.87)	mn: 1.23 (0.31; 2.15)	mn: 1.15 (1.02; 1.28)	**<0.001**
Median arterial pressure z-score for age (first 24 h min)	md: −0.18 (−0.23; 0.06)	md: −0.72	md: 0.04 (−0.11; 0.09)	**<0.001**
"(first 24 h mean)	mn: 0.15 (0.11; 0.18)	mn: 0.15 (−0.02; 0.32)	mn: 0.41 (0.35; 0.47)	**<0.001**

Table A1. *Cont.*

	Cluster 1	Cluster 2	Cluster 3	*p* Value
Variables used for clustering	(Mean/Median (CI95%))	(Mean/Median (CI95%))	(Mean/Median (CI95%))	
"(first 24 h max)	mn: 0.54 (0.47; 0.62)	mn: 0.78 (0.06; 1.5)	mn: 0.85 (0.76; 0.94)	**<0.001**
Systolic arterial pressure z-score for age (first 24 h min)	mn: −0.21 (−0.28; −0.14)	mn: −0.35 (−0.58; −0.12)	mn: −0.01 (−0.07; 0.06)	**<0.001**
"(first 24 h mean)	mn: 0.1 (0.05; 0.14)	mn: −0.07 (−0.16; 0.02)	mn: 0.32 (0.27; 0.36)	**<0.001**
"(first 24 h max)	md: 0.37 (0.35; 0.58)	md: 0.32	md: 0.61 (0.56; 0.69)	**<0.001**
First 24 h diuresis in mL/kg/h	mn: 1.85 (1.16; 2.54)	mn: 3.4 (0; 7.27)	mn: 1.51 (0.68; 2.33)	0.90
Temperature in °C (first 24 h min)	mn: 36.74 (36.46; 37.02)	mn: 34.82 (34.19; 35.46)	mn: 36.7 (36.22; 37.17)	0.50
"(first 24 h mean)	mn: 37.02 (36.75; 37.28)	mn: 35.57 (35.15; 35.98)	mn: 37.15 (36.72; 37.58)	0.97
"(first 24 h max)	mn: 37.31 (36.96; 37.67)	mn: 36.18 (35.65; 36.7)	mn: 37.7 (37.18; 38.21)	0.43
Oxygen saturation in % (first 24 h min)	md: 94 (93; 95)	md: 79 (61; 89)	md: 95.5 (94; 96)	0.002
"(first 24 h mean)	md: 97.88 (96.94; 98.48)	md: 94 (91.33; 98.71)	md: 98.27 (97.38; 98.86)	0.02
Venous oxygen saturation in % (first 24 h min)	mn: 56.12 (50.16; 62.08)	mn: 34.53 (24.92; 44.14)	mn: 62.96 (57.16; 68.76)	0.30
"(first 24 h max)	mn: 77.24 (71.62; 82.85)	mn: 71.73 (59.86; 83.61)	mn: 78.6 (74.59; 82.61)	0.81
Carboxyhemoglobin in % (first 24 h max)	mn: 2.3 (1.89; 2.71)	mn: 2.49 (2.03; 2.94)	mn: 2.55 (2.1; 3)	0.38
Inspirited oxygen fraction in % (first 24 h max)	mn: 57.78 (46.9; 68.65)	mn: 88.33 (76.73; 99.93)	mn: 46.53 (31.11; 61.96)	0.81
"(first 24 h mean)	md: 36.19 (33.1; 55.62)	md: 55.59 (40; 69.76)	md: 26 (26; NA)	**<0.001**
Capillary glucose in mg/dL (first 24 h min)	mn: 100.11 (90.91; 109.31)	mn: 75.53 (53.84; 97.22)	mn: 111.77 (96.74; 126.81)	0.39
"(first 24 h max)	md: 156.5 (126; 224)	md: 228 (199; 310)	md: 160.5 (135; 253)	0.01
Calcium ion in mg/dL (first 24 h min)	md: 4.6 (4.5; 4.9)	md: 3.6 (3.2; 4)	md: 4.54 (4.4; 4.7)	**<0.001**
Creatinine in mg/dL (first 24 h max)	mn: 3.18 (0; 7.05)	mn: 3.92 (0; 10.08)	mn: 3.98 (0; 7.98)	0.77
Phosphate in mg/dL (first 24 h max)	md: 4.8 (4.5; 6.7)	md: 8 (5.8; 10.7)	md: 4.25 (3.9; 5.1)	**<0.001**
Reactive C protein in mg/L (first 24 h max)	md: 55.9 (50.2; 198.09)	md: 11.4 (7.73; 38.94)	md: 163.6 (109.65; 265.62)	**<0.001**
Lymphocytes · 1000/µL (first 24 h min)	md: 1.4 (0.92; 4.6)	md: 0.9 (0.6; 2.33)	md: 0.21 (0.05; 0.8)	**<0.001**
"(first 24 h max)	md: 3.24 (2.88; 25)	md: 4.42 (1.5; 21)	md: 0.24 (0.1; 1.2)	**<0.001**
Neutrophils · 1000/µL (first 24 h min)	md: 7.25 (6.5; 21.88)	md: 2.65 (2; 4.62)	md: 1.43 (0.05; 6.73)	**<0.001**
"(first 24 h max)	mn: 13.65 (9.61; 17.69)	mn: 5.53 (3.83; 7.23)	mn: 6.01 (2.77; 9.25)	0.002
Mechanical ventilation	mn: 0.78 (0.66; 0.91)	mn: 1	mn: 0.58 (0.37; 0.78)	0.005
Hemodiafiltration	p: 0.02 (0; 0.07)	p: 0.28 (0.05; 0.51)	p: 0.04 (0; 0.12)	0.002
ECMO	p: 0.02 (0; 0.07)	p: 0.83 (0.64; 1.02)	p: 0	**<0.001**
Thermic blanket	p: 0.11 (0.02; 0.2)	p: 0.06 (0; 0.17)	p: 0	0.20
Non-clustering variables	(mean/median (CI95%))	(mean/median (CI95%))	(mean/median (CI95%))	
EtCO2 in Torr (first 24 h min)	md: 30 (40; NA)	md: 9 (8; NA)	md: 31	**<0.001**
"(first 24 h mean)	mn: 42.56 (35.27; 49.86)	mn: 22.33 (16.54; 28.12)	mn: 35.19 (25.63; 44.74)	0.02
"(first 24 h max)	mn: 51.35 (43.8; 58.9)	mn: 33.08 (22.85; 43.32)	mn: 43.25 (32.67; 53.83)	0.04
End expiratory pressure (first 24 h mean)	md: 5.5	md: 6.97	md: 5.14	0.13
"(first 24 h max)	md: 6	md: 9	md: 6	0.10
Cerebral NIRS in % (first 24 h min)	mn: 49.67 (39.43; 59.9)	mn: 47.25 (35.27; 59.23)	mn: 61.25 (45.11; 77.39)	0.44
"(first 24 h mean)	mn: 59.59 (47.6; 71.59)	mn: 66.11 (59.64; 72.58)	mn: 70.73 (64.26; 77.2)	0.12
Troponin in ng/L (first 24 h max)	md: 249 (35; 1653)	md: 2732 (533; 149197.5)	md: 29 (20; 35)	0.01
Cardiogenic shock	p: 0.2 (0.08; 0.31)	p: 0.72 (0.49; 0.95)	p: 0.08 (0; 0.19)	**<0.001**
Inflammatory shock	p: 0.52 (0.37; 0.67)	p: 0.17 (0; 0.36)	p: 0.81 (0.65; 0.97)	**<0.001**
Septic shock	p: 0.5 (0.35; 0.65)	p: 0.17 (0; 0.36)	p: 0.77 (0.6; 0.94)	**<0.001**
Anaphylactic shock	p: 0.04 (0; 0.1)	p: 0	p: 0	0.38
Hypovolemic shock	p: 0.15 (0.04; 0.26)	p: 0.28 (0.05; 0.51)	p: 0.12 (0; 0.25)	0.34
Hypovolemic secondary to traumatism	p: 0.07 (0; 0.14)	p: 0	p: 0.04 (0; 0.12)	0.52

Table A1. *Cont.*

Non-clustering variables	Cluster 1 (Mean/Median (CI95%))	Cluster 2 (Mean/Median (CI95%))	Cluster 3 (Mean/Median (CI95%))	*p* Value
Hypovolemic secondary to surgery	p: 0.02 (0; 0.07)	p: 0.06 (0; 0.17)	p: 0.04 (0; 0.12)	0.78
Non-specified shock	p: 0.2 (0.08; 0.31)	p: 0	p: 0.04 (0; 0.12)	**0.03**
Cardiac surgery	mn: 0.2 (0.08; 0.31)	mn: 0.72 (0.49; 0.95)	mn: 0.08 (0; 0.19)	**<0.001**
Oncologic patients	mn: 0.04 (0; 0.1)	mn: 0.06 (0; 0.17)	mn: 0.31 (0.12; 0.5)	**0.003**
Length of stay in days	md: 6 (3; 9)	md: 11 (4; 29)	md: 4 (2; 7)	**0.02**
Exitus	p: 0.07 (0; 0.14)	p: 0.33 (0.09; 0.57)	p: 0.04 (0; 0.12)	**0.004**

References

1. Nichols, D.G.; Shaffner, D.H.; Argent, A.C. *Rogers' Textbook of Pediatric Intensive Care*; Wolters Kluwer: Philadelphia, PA, USA, 2016.
2. Kliegman, R.; Marcdante, K.J. *Nelson Essentials of Pediatrics*; Elsevier: Amsterdam, The Netherlands, 2019.
3. Standl, T.; Annecke, T.; Cascorbi, I.; Heller, A.R.; Sabashnikov, A.; Teske, W. The Nomenclature, Definition and Distinction of Types of Shock. *Dtsch. Arztebl. Int.* **2018**, *115*, 757–768. [CrossRef] [PubMed]
4. Kislitsina, O.N.; Rich, J.D.; Wilcox, J.E.; Pham, D.T.; Churyla, A.; Vorovich, E.B.; Ghafourian, K.; Yancy, C.W. Shock—Classification and Pathophysiological Principles of Therapeutics. *Curr. Cardiol. Rev.* **2019**, *15*, 102–113. [CrossRef] [PubMed]
5. Peters, M.J.; Shipley, R. Clinical Classification of Cold and Warm Shock: Is There a Signal in the Noise? *Pediatr. Crit. Care Med.* **2020**, *21*, 1085–1087. [CrossRef] [PubMed]
6. Tibby, S.M.; Hatherill, M.; Marsh, M.J.; Murdoch, I.A. Clinicians' abilities to estimate cardiac index in ventilated children and infants. *Arch. Dis. Child.* **1997**, *77*, 516–518. [CrossRef] [PubMed]
7. Benke, K.; Benke, G. Artificial Intelligence and Big Data in Public Health. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2796. [CrossRef] [PubMed]
8. Growth Charts—Data Table of Infant Weight-for-Age Charts. Centers for Disease Control and Prevention. Available online: https://www.cdc.gov/growthcharts/html_charts/wtageinf.htm (accessed on 9 August 2022).
9. Growth Charts—Data Table of Weight-for-Age Charts. Centers for Disease Control and Prevention. Available online: https://www.cdc.gov/growthcharts/html_charts/wtage.htm (accessed on 9 August 2022).
10. Flynn, J.T.; Kaelber, D.C.; Baker-Smith, C.M.; Blowey, D.; Carroll, A.E.; Daniels, S.R.; De Ferranti, S.D.; Dionne, J.M.; Falkner, B.; Flinn, S.K.; et al. Clinical Practice Guideline for Screening and Management of High Blood Pressure in Children and Adolescents. *Pediatrics* **2017**, *140*, e20171904. [CrossRef] [PubMed]
11. Fleming, S.; Thompson, M.; Stevens, R.; Heneghan, C.; Plüddemann, A.; Maconochie, I.; Tarassenko, L.; Mant, D. Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: A systematic review of observational studies. *Lancet* **2011**, *377*, 1011. [CrossRef]
12. Likas, A.; Vlassis, N.; Verbeek, J.J. The global K-means clustering algorithm. *Pattern Recognit.* **2003**, *36*, 451–461. [CrossRef]
13. Dokmanic, I.; Parhizkar, R.; Ranieri, J.; Vetterli, M. Euclidean distance matrices: Essential theory, algorithms, and applications. *IEEE Signal Processing Mag.* **2015**, *32*, 12–30. [CrossRef]
14. Mohamad, I.B.; Usman, D. Standardization and its effects on K-means clustering algorithm. *Res. J. Appl. Sci. Eng. Technol.* **2013**, *6*, 3299–3303. [CrossRef]
15. Jia, Z.; Lu, X.; Duan, H.; Li, H. Using the distance between sets of hierarchical taxonomic clinical concepts to measure patient similarity. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 91. [CrossRef]
16. Marlin, B.M.; Kale, D.C.; Khemani, R.G.; Wetzel, R.C. Unsupervised pattern discovery in electronic health care data using probabilistic clustering models. In Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, Miami, FL, USA, 28–30 January 2012; ACM: New York, NY, USA, 2012; pp. 389–398.
17. Williams, J.B.; Ghosh, D.; Wetzel, R.C. Applying Machine Learning to Pediatric Critical Care Data. *Pediatr. Crit. Care Med.* **2018**, *19*, 599–608. [CrossRef]
18. Hug, C.W.; Clifford, G.D.; Reisner, A.T. Clinician blood pressure documentation of stable intensive care patients: An intelligent archiving agent has a higher association with future hypotension. *Crit. Care Med.* **2011**, *39*, 1006–1014. [CrossRef] [PubMed]
19. Wetzel, R.C. First get the data—Then do the science. *Pediatr. Crit. Care Med.* **2018**, *19*, 382–383. [CrossRef] [PubMed]