BioData Mining

RESEARCH

Open Access

Temporal phenotyping and prognostic stratification of patients with sepsis through longitudinal clustering

Patrizia Ribino[1*], Maria Mannone[1,2], Claudia Di Napoli[3], Giovanni Paragliola[3], Davide Chicco[4,5,6*] and Francesca Gasparini[4,5]

*Correspondence: patrizia.ribino@icar.cnr.it; davide.chicco@unimib.it

[1] Istituto di Calcolo e Reti ad Alte Prestazioni, Consiglio Nazionale delle Ricerche (CNR), Palermo, Italy
[2] Institute of Physics and Astronomy, Universität Potsdam, Potsdam, Germany
[3] Istituto di Calcolo e Reti ad Alte Prestazioni, Consiglio Nazionale delle Ricerche (CNR), Naples, Italy
[4] Dipartimento di Informatica Sistemistica e Comunicazione, Università di Milano-Bicocca, Milan, Italy
[5] NeuroMi, Milan Centre for Neuroscience, Milan, Italy
[6] Institute for Health Policy Management and Evaluation, University of Toronto, Toronto, Ontario, Canada

Abstract

Sepsis is a critical medical condition characterized by a highly variable and rapidly evolving clinical course, often necessitating early intervention and tailored treatment plans to improve patient outcomes. Due to its complexity and heterogeneity, understanding the progression of sepsis across different patient populations remains a significant challenge. In this study, we exploit a sophisticated analytical framework based on k-means multivariate longitudinal clustering to capture the diverse trajectories of sepsis. We do so by analyzing multiple clinical parameters tracked over time, providing a nuanced view of disease progression. By incorporating Dynamic Time Warping (DTW) as the distance metric, the proposed method effectively accounts for temporal misalignments and variability in the rate of disease progression, an essential capability given the unpredictable and heterogeneous nature of sepsis. This integration enhances the model's ability to detect distinct temporal patterns and phenotypic subgroups that may remain undetected using conventional analytical approaches. By leveraging sepsis-related electronic health records (EHRs), which provide rich time-series data on laboratory results along with patient demographics and underlying health conditions, the proposed method reveals distinct sepsis phenotypes that reflect variations in disease progression. We perform several experiments varying the number of clusters and clinical variable combinations, evaluating the clustering performances using Silhouette score, Caliski-Harabasz Index, and Davies-Bouldin Index, as reference quality metrics. Our results confirm the prognostic role of the Thrombin-Antigen complex and the Prothrombin Time–International Normalized Ratio for septic patients. Furthermore, to evaluate the relevance of subjects' stratification, the Adjusted Rand Index metric is used to quantify the survival prediction capability of our longitudinal clustering method, considering the 28-day death feature as the target variable. The same metric demonstrates that our proposal outperforms other longitudinal clustering algorithms available in the literature.

Keywords: Clustering, Unsupervised machine learning, Electronic health records, Longitudinal clustering, Patient trajectories, Sepsis, Intensive care unit

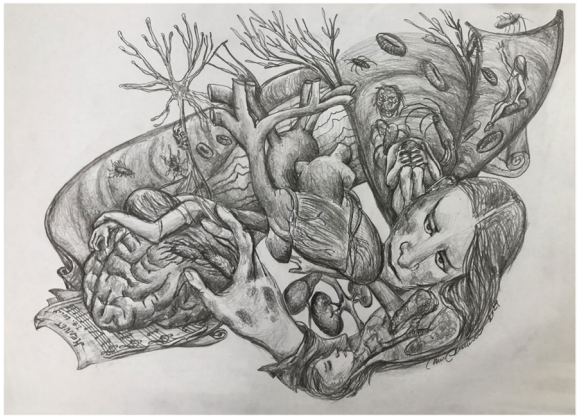

Fig. 1 Sepsis. Drawing by Maria Mannone

Introduction

An infected wound, a gangrene, a refused amputation, and ultimately a death for sepsis: this is the tragic fate of the French composer Jean-Baptiste Lully (1632-1687)[1]. Also today, sepsis is still causing silent and rapid deaths. It is a complex, life-threatening condition caused by the body's overwhelming response to infection, often leading to multi-organ failure or septic shock [1]. Sepsis can arise from several types of infection, including bacterial, viral, fungal, or parasitic infections. Familiar sources of infection that lead to sepsis include Pneumonia, Urinary tract infections (UTIs), Abdominal infections, Bloodstream infections (bacteremia), and surgical site infections. In sepsis, the body's immune system triggers an excessive and harmful inflammatory response to fight infection. This can result in (i) Systemic Inflammatory Response Syndrome (SIRS) [2], namely a widespread release of inflammatory molecules that cause damage to blood vessels, leading to fluid leakage, edema, and reduced blood flow to vital organs, and (ii) coagulopathy [3] caused by the activation of the clotting system that leads to the formation of small blood clots (microthrombi) in the microcirculation, which can further reduce blood flow and contribute to organ dysfunction. This can eventually lead to Disseminated Intravascular Coagulation (DIC), characterized by excessive clotting and bleeding. Finally, Multiple Organ Dysfunction Syndrome [4] due to the reduced blood flow and tissue damage caused by inflammation and micro clot formation can compromise the function of major organs, including the heart, lungs, liver, kidneys, and brain. We can pictorially summarize the multi-organ damage occurring in sepsis through the representation of Fig. 1.

Today, we can investigate sepsis via data analysis and machine learning strategies. In this endeavor, electronic health records (EHRs) play a pivotal role in identifying, managing, and studying sepsis in clinical practice. Sepsis prediction models integrated into EHRs can exploit real-time patient data, and advanced machine learning (ML) techniques could enable early identification of patients at risk of developing sepsis, often

[1] At that time, to mark tempo for orchestral conducting, a large baton was rhythmically hit on the ground. Once, when Lully was conducting, he injured his feet with the baton. The wound then resulted in gangrene. Since Lully was an able dancer, he refused the amputation of his leg, but the following sepsis ultimately led to his death.

before clinical signs become evident. These models have the potential to significantly improve sepsis outcomes by allowing for earlier interventions, such as antibiotic administration and fluid resuscitation, reducing the risk of severe complications or death. In [5], the authors developed a prediction model to evaluate the probability transition between three different disease states to estimate the daily evolution of disease severity during sepsis. XGBoost and LightGBM have been applied in [6] to predict early sepsis six hours in advance. The authors in [7] used an LSTM network for the early detection of septic shock, showing that the proposed method allows one to detect patients up to 20 hours earlier.

Among the various ML techniques, unsupervised longitudinal clustering holds considerable potential for application within the context of electronic health records. The development of an EHR's longitudinal clustering allows for the grouping of patients based on their health data collected over time. Indeed, EHRs are longitudinal by nature, meaning they capture patients' medical history, diagnoses, treatments, and outcomes over extended periods. Applying longitudinal clustering to these records allows for analyzing patterns or trajectories within the data and for identifying subgroups of patients who exhibit similar health trends or progression of conditions.

Recently, longitudinal clustering has emerged as a crucial approach for capturing disease progression dynamics in clinical research. For example, Amaral et al. [8] and Ramamoorthy et al. [9] applied temporal clustering to amyotrophic lateral sclerosis cohorts, revealing meaningful subgroups linked to progression trajectories. Similarly, recent work in Alzheimer's disease demonstrated that multivariate longitudinal clustering can uncover neuropsychological predictors of dementia progression, further highlighting the clinical utility of time-aware patient stratification [10].

In the context of sepsis, besides several works that apply static k-means to investigate phenotypic subgroups, including [11–13], several traditional unsupervised clustering models have been applied to group septic patients, analyzing their data at specific time points. For instance, spectral clustering has been applied in [14] to find four groups of patients and study the trajectories of physiological variables associated with their risk-score clusters. Hierarchical clustering in [15] has been applied to identify clusters of patients according to clinical and biological characteristics collected at patients' admission. On the other hand, to the best of our knowledge, only a few works have adopted unsupervised longitudinal machine learning approaches. Among them, the authors in [16] and in [17] clustered groups of patients in heterogeneous medical conditions, using longitudinal k-means [18], and tracking longitudinal biomarkers, to understand the progression of Acute Kidney Injury and sepsis and their impact on mortality in patients with burns. Several studies applied supervised ML approaches to datasets derived from static (single visit) EHRs of patients with sepsis or septic shock [19–21], but we found them unrelated to multi-visit longitudinal datasets.

In this study, we propose the adoption of our k-means multivariate longitudinal clustering [10] for stratifying patients based on the similarity of their temporal profiles over the course of sepsis. We adopt Dynamic Time Warping (DTW) as a method for measuring the similarity between two time-series data points even if their trajectories differ in timing. The goal is to group patients based on the similarity of their sepsis-related trajectories rather than static measurements. For sepsis, DTW can align different patients'

clinical variables, even if they deteriorate or recover at different rates. Patients in the same cluster it is expected to have similar sepsis progression patterns across multiple variables. In contrast, those in different clusters might represent distinct clinical phenotypes of sepsis (for example, fast versus slow progression).

Within this context, the main objective of this paper is to identify the clinical variables that better stratify subjects in terms of sepsis progression. To reach this goal, we perform a longitudinal study applying our longitudinal clustering algorithm [10] on a publicly available sepsi dataset [22].

Our study addresses two different aspects:

- Identifying the configuration in terms of the number of clusters and clinical variables that better group the subjects. To this end and to provide more explainable results, we perform several experiments varying the number of clusters and combinations of clinical variables, evaluating them with clustering quality metrics, such as Silhouette score [23], Caliski-Harabasz Index [24], and Davies-Bouldin Index [25].
- Evaluating the relevance of subjects' stratification obtained by our longitudinal clustering algorithm in the real clinical scenario. Such a relevance is assessed by the survival prediction capability of our algorithm, measured with the Adjusted Rand Index metric [26], considering the 28-day death feature as the target variable. We also compare our results with those obtained by applying other longitudinal clustering algorithms available in the literature by adopting the same metric.

The rest of the paper is organized as follows. Materials and methods used in this work are reported in Dataset and Multivariate longitudinal cluster analysis sections. Evaluation of the proposed longitudinal clustering as well as a comparison with other longitudinal approaches are reported in Clinical variables identification and assessment section. Results are discussed in Discussion section, and conclusions are drawn in Conclusions section.

Dataset

The dataset we analyze in this study was collected at the Jichi Medical University Hospital in Shimotsuke, Tochigi, Japan and was derived from data of patients admitted to the intensive care unit (ICU) of the university hospital between April 2014 and September 2016 [22]. The dataset was described by the data curators and released publicly in a 2018 study [22], following the FAIR (findability, accessibility, interoperability, and reusability) principles [27]. The original dataset contains data of ICU patients and includes both static (single-visit or status) features and longitudinal (multi-visit) variables. The static features include the sex of the patient, the source of sepsis (i.e., pulmonary, abdominal, urinary tract, soft tissue, bloodstream, or other), and one or more comorbidities that could be present for a patient in addition to sepsis (ischemic heart disease, congestive heart failure, arrhythmia, chronic obstructive pulmonary disease, chronic kidney disease, and/or cardiovascular diseases).

The dataset includes other static variables, such as APACHE II (Acute Physiology and Chronic Health Evaluation) and the SOFA (Sequential Organ Failure Assessment) scores. APACHE II is commonly used in the ICU to predict outcomes in critically ill

patients, including those with sepsis. It considers a variety of physiological parameters (such as temperature, blood pressure, heart rate, oxygenation) and chronic health conditions to estimate the risk of mortality. The SOFA score is used to assess the dysfunction of several organ systems in the body and to predict mortality in a population of sepsis patients. Moreover, the dataset also includes the 28-day death (also known as 28-day mortality), i.e., the occurrence of death, from any cause, within 28 days after the septic episode. Additional information about this dataset can be found in the original study [22].

Table 1 summarizes the demographic and clinical characteristics of the study population ($n = 204$) upon ICU admission, including age, sex, illness severity scores (APACHE II and SOFA), length of ICU stay, 28-day mortality, and sepsis source distribution. Notably, the original dataset reveals substantial variability in age (range 19–101 years), illness severity (APACHE II scores ranging from 7 to 53 and SOFA from 2 to 19), and prognosis, suggesting a heterogeneous patient population. The distribution of sepsis sources also varies, with abdominal and pulmonary infections being the most common.

Table 1 Demographics and clinical characteristics of patients at ICU admission for the original dataset

Static Features	**Dataset ($n = 204$)**
Demographics	
Age (years)	
mean ± SD	68.3 ± 14.5
[min, max]	[19, 101]
Gender	
Male n°(%)	117 (57.4%)
Female n°(%)	87 (42.6%)
Illness severity	
APACHE II	
mean ± SD	25.3 ± 8.1
[min, max]	[7, 53]
SOFA	
mean ± SD	8.2 ± 3.4
[min, max]	[2, 19]
Source of sepsis	
Pulmonary n°(%)	50 (24.5%)
Abdominal n°(%)	101 (49.5%)
Urinary n°(%)	11 (5.4%)
Soft tissue n°(%)	26 (12.7%)
Blood n°(%)	2 (0.9%)
other n°(%)	14 (6.8%)
Prognosis	
ICU days	
mean ± SD	10.6 ± 7.7
[min, max]	[2, 58]
28-day death	
survived n°(%)	182 (89.2%)
death n°(%)	22 (10.8%)

In our longitudinal study, only clinical variables that have a temporal dynamic are considered. In addition, clinical variables strictly correlated to the prognosis and severity of sepsis are not considered to avoid bias in the results. Hence, for each patient, the following hemostatic and thrombopoietic clinical variables are considered, each one collected over a period of seven ICU days, after the sepsis diagnosis on ICU admission:

- The Thrombin-Antithrombin (TAT) complex is a biomarker that indicates the activation of the coagulation system. The TAT complex is often measured in clinical settings to assess the level of thrombin generation and overall coagulation activity.
- The Prothrombin Time–International Normalized Ratio (INR) is a laboratory measurement used to assess the time it takes for blood to clot. It is a standardized Prothrombin Time (PT) version, allowing consistent results across different labs [28].
- The Fibrin Degradation Products (FDP) are fragments resulting from fibrin breakdown in blood clots. They are a crucial marker of fibrinolysis.
- The Absolute Immature Platelet Count (AIPC) measures the immature platelets in the bloodstream. Immature platelets are newly produced by the bone marrow and released into circulation. They reflect the bone marrow's platelet production rate and are often used as a marker of thrombopoiesis (platelet formation).
- Protein C (also called Anticoagulant Protein C, APC) is a vital anticoagulant protein in the body, playing a key role in regulating blood clotting and preventing excessive clot formation [29].
- Platelet Count (PlC) is a crucial component of a Complete Blood Count (CBC) that measures the number of platelets in a given blood volume. Platelets are small cell fragments that play a key role in blood clotting and wound repair by aggregating at the site of vascular injury and forming a clot.

Table 2 reports the values of longitudinal clinical variables at ICU admission, showing wide inter-individual differences across all biomarkers. For example, TAT levels range from 0.7 to 290.8 ng/mL, and FDP values span from 2.8 to 256.2 μg/mL, reflecting diverse coagulation profiles. Such variability further supports the hypothesis of underlying heterogeneity of patients and may warrant subgroup analysis.

Multivariate longitudinal cluster analysis

The values of clinical variables over time for each patient can be seen as trajectories in a multidimensional space. We intend to cluster patients' trajectories according to their similarity to identify prognostic patient stratification of medical interest. To this aim, we developed a k-means longitudinal clustering with soft Dynamic Time Warping (soft-DTW) to discover patterns of joint trajectory in multivariate time series [10]. The pipeline of the proposed approach is sketched in Fig. 2. Given a group of septic patients characterized by a set of features collected over a period of time, a z-score normalization is performed in order to standardize the temporal trajectories across patients by removing scale-related biases to allow for data comparison.

Variable time alignment was not required in this case, as all timepoints were consistently present across patients. Regarding missing data, we note that no imputation or data augmentation methods were necessary, as each patient had complete blood

Table 2 Hemostatic and thrombopoietic parameters of patients at ICU admission for the original dataset

Longitudinal Features	**Dataset** ($n = 204$)
Platelet Count	
mean ± SD	14.9 ± 7.4
[min, max]	[0.6, 43.4]
AIPC	
mean ± SD	7.3 ± 4.3
[min, max]	[1.1, 28.6]
INR	
mean ± SD	1.5 ± 0.5
[min, max]	[0.9, 4.3]
FDP	
mean ± SD	27.4 ± 28.1
[min, max]	[2.8, 256.2]
TAT	
mean ± SD	21.6 ± 28.2
[min, max]	[0.7, 290.8]
Protein C	
mean ± SD	48.7 ± 22.2
[min, max]	[11.7, 169.4]

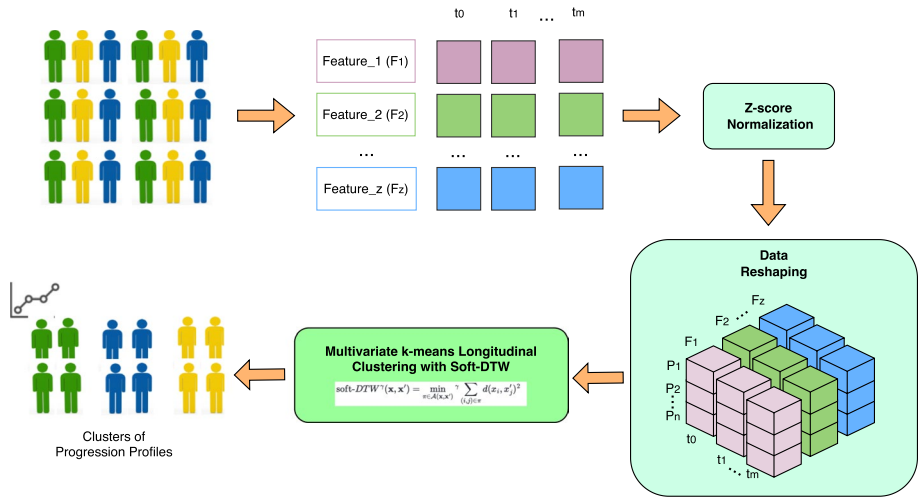

Fig. 2 Pipeline for stratifying septic patients according to hemostatic and thrombopoietic parameters collected over seven days

test records for all seven days, except in cases of early mortality. In such cases, the absence of data simply reflects the clinical outcome rather than technical missingness.

Data are then reshaped into a three-dimensional matrix of (samples, time, features) as input to our multivariate k-means longitudinal clustering method, whose key concepts are here summarized.

Let us consider a set of p individuals. For each individual i, a set of n clinical variables (features) are measured at each visit t, defining the individual's health state at time t, as follows:

$$F_i(t) = \{f_{i1}(t), f_{i2}(t), \ldots f_{in}(t)\} \tag{1}$$

The joint trajectory for individual i for the time window $t \in [0, m]$ is defined as follows:

$$JointTraj_i = (F_i(0), F_i(1), \ldots F_i(m)) \tag{2}$$

where $F_i(0), F_i(1), \ldots F_i(m)$ are the health states of individual i at $t = 0$, i.e. at the first visit, and at the successive visits.

The pseudocode of our multivariate longitudinal clustering is reported in Algorithm 1 box. It is based on the well-known k-means algorithm [30] that partitions a dataset comprising p observations into k internally homogeneous clusters. It begins by initializing the centroids, typically randomly selected within the feature space. Each observation is then assigned to a cluster based on its proximity to the nearest centroid. To achieve an optimal clustering configuration, k-means alternates between two iterative phases analogous to the Expectation-Maximization (EM) framework [31]. In the Expectation (E) step, the algorithm computes the distance between each observation and the current cluster centroids. In the Maximization (M) step, observations are reassigned to the cluster whose centroid is closest. These steps are repeated iteratively until convergence, typically when cluster assignments stabilize.

Algorithm 1 Multivariate k-means for longitudinal clustering

Data: $Data = \{JointTraj_i | i \in [1, p]\}$
Result: $Centroids : \{C_k\} = \{c_{kj}(t) | j \in [1, w], w \leq n, t \in [1, m]\}, \forall k \in [1, Nclust]$
$Initial\_Features \leftarrow \{f_1, f_2, \ldots f_n\}$;
$k \leftarrow 2$;
while $k < Nclust$ **do**
  $iteration \leftarrow 0$;
  $Centroids \leftarrow \varnothing$;
  $Clusters \leftarrow \varnothing$;
  $Data_{current} \leftarrow get(Data, Initial\_Features)$;
  /*Cluster Initialization*/
  for $c = 1, ..., k$ **do**
    $C_c \leftarrow RandomChoose(Data_{current})$;
    $Clust_c \leftarrow RandomAssign(Data_{current}, C_c)$;
  $Centroids \leftarrow \{C_1, ..., C_k\}$;
  $Clusters \leftarrow \{Clust_1, ..., Clust_k\}$;
  /*Cluster Formation*/
  repeat
    $iteration \leftarrow iteration + 1$;
    for $i = 1, ..., p$ **do**
      $List\_Dist \leftarrow \varnothing$
      for $c = 1, ..., k$ **do**
        $dist(c) \leftarrow SoftDTWDist(JointTraj_i, C_c)$;
      $List\_Dist \leftarrow \{dist(1), ..., dist(k)\}$;
      $closestC \leftarrow indexOf(min(List\_Dist))$;
      $Clusters \leftarrow assign(JointTraj_i, closestC)$;
    $Centroids \leftarrow SoftDTWBarycenter(Clusters)$;
    /*Clustering Stop Condition Evaluation */
    for $c = 1, ..., k$ **do**
      $\delta(c) \leftarrow SoftDTWDist(C_{c_{iter-1}}, C_{c_{iter}})$;
    $\Delta \leftarrow \{\delta(1), ..., \delta(k)\}$
    if $\Delta < Threshold$ **then**
      stop = True;
  until $iteration \neq maxIteration$ OR $stop = True$;
  $EvaluateMetrics(Clusters)$;

Data: $Data = \{JointTraj_i | i \in [1, p]\}$
Result: $Centroids : \{C_k\} = \{c_{kj}(t) | j \in [1, w], w \leq n, t \in [1, m]\}, \forall k \in [1, Nclust]$
$Initial\_Features \leftarrow \{f_1, f_2, \dots f_n\}$;
$k \leftarrow 2$;
while $k < Nclust$ **do**
  $iteration \leftarrow 0$;
  $Centroids \leftarrow \varnothing$;
  $Clusters \leftarrow \varnothing$;
  $Data_{current} \leftarrow get(Data, Initial\_Features)$;
  /*Cluster Initialization*/
  for $c = 1, ..., k$ **do**
    $C_c \leftarrow RandomChoose(Data_{current})$;
    $Clust_c \leftarrow RandomAssign(Data_{current}, C_c)$;
  $Centroids \leftarrow \{C_1, ..., C_k\}$;
  $Clusters \leftarrow \{Clust_1, ..., Clust_k\}$;
  /*Cluster Formation*/
  repeat
    $iteration \leftarrow iteration + 1$;
    for $i = 1, ..., p$ **do**
      $List\_Dist \leftarrow \varnothing$
      for $c = 1, ..., k$ **do**
        $dist(c) \leftarrow SoftDTWDist(JointTraj_i, C_c)$;
      $List\_Dist \leftarrow \{dist(1), ..., dist(k)\}$;
      $closestC \leftarrow indexOf(min(List\_Dist))$;
      $Clusters \leftarrow assign(JointTraj_i, closestC)$;
    $Centroids \leftarrow SoftDTWBarycenter(Clusters)$;
    /*Clustering Stop Condition Evaluation */
    for $c = 1, ..., k$ **do**
      $\delta(c) \leftarrow SoftDTWDist(C_{c_{iter-1}}, C_{c_{iter}})$;
    $\Delta \leftarrow \{\delta(1), ..., \delta(k)\}$
    if $\Delta < Threshold$ **then**
      stop = True;
  until $iteration \neq maxIteration \ OR \ stop = True$;
  $EvaluateMetrics(Clusters)$;

Data: $Data = \{JointTraj_i | i \in [1, p]\}$
Result: $Centroids : \{C_k\} = \{c_{kj}(t) | j \in [1, w], w \leq n, t \in [1, m]\}, \forall k \in [1, Nclust]$
$Initial\_Features \leftarrow \{f_1, f_2, \dots f_n\}$;
$k \leftarrow 2$;
while $k < Nclust$ **do**
  $iteration \leftarrow 0$;
  $Centroids \leftarrow \varnothing$;
  $Clusters \leftarrow \varnothing$;
  $Data_{current} \leftarrow get(Data, Initial\_Features)$;
  /*Cluster Initialization*/
  for $c = 1, ..., k$ **do**
    $C_c \leftarrow RandomChoose(Data_{current})$;
    $Clust_c \leftarrow RandomAssign(Data_{current}, C_c)$;
  $Centroids \leftarrow \{C_1, ..., C_k\}$;
  $Clusters \leftarrow \{Clust_1, ..., Clust_k\}$;
  /*Cluster Formation*/
  repeat
    $iteration \leftarrow iteration + 1$;
    for $i = 1, ..., p$ **do**
      $List\_Dist \leftarrow \varnothing$
      for $c = 1, ..., k$ **do**
        $dist(c) \leftarrow SoftDTWDist(JointTraj_i, C_c)$;
      $List\_Dist \leftarrow \{dist(1), ..., dist(k)\}$;
      $closestC \leftarrow indexOf(min(List\_Dist))$;
      $Clusters \leftarrow assign(JointTraj_i, closestC)$;
    $Centroids \leftarrow SoftDTWBarycenter(Clusters)$;
    /*Clustering Stop Condition Evaluation */
    for $c = 1, ..., k$ **do**
      $\delta(c) \leftarrow SoftDTWDist(C_{c_{iter-1}}, C_{c_{iter}})$;
    $\Delta \leftarrow \{\delta(1), ..., \delta(k)\}$
    if $\Delta < Threshold$ **then**
      stop = True;
  until $iteration \neq maxIteration \ OR \ stop = True$;
  $EvaluateMetrics(Clusters)$;

In the context of longitudinal data analysis, the concept of a cluster centroid is extended to represent the average trajectory of all individuals within a given cluster. That is, the centroid corresponds to a temporal pattern derived from the multivariate time series of all members in the group. For a given individual i, the most appropriate cluster C is defined as the one whose trajectory minimizes the temporal distance between the individual's trajectory and the cluster centroid. Such temporal distance is computed using Dynamic Time Warping (DTW) distance [32]. Different from the Euclidean distance, DTW detects a similarity of shape between two curves. Considering as (discrete) curves the polylines joining the points of two time-series, respectively, the shift between them can occur over time. Considering time as the dependent variable, DTW finds the optimal correspondence between the two considered time series [33]. Here, we adopt soft-DTW, with a smoothed and differentiable cost function. In soft-DTW, all possible temporal alignments, rather than the optimal one (as for DTW), are considered [34]. The smoothing degree is user-adjusted via the parameter γ. Unlike traditional DTW, which gives a binary "match" or "mismatch" between points in two time-series sequences, soft-DTW assigns a soft cost for the warping path. This means that it allows for more flexibility in aligning time-series data that may have noise, outliers, or slight misalignments [35]. The integration of soft-DTW in k-means for longitudinal clustering entails moving from traditional centroid evaluation using Euclidean distance to computing barycenters

that are time series themselves, constructed to minimise their soft-DTW distance to the other series.

Cluster evaluation

In our work, we consider three internal assessment metrics and an external assessment metric. The following internal assessment metrics have been considered to evaluate the performance of the longitudinal clustering. The Silhouette coefficient (S) [23], which measures the quality of cluster placement for each individual i, is computed as:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}, \tag{3}$$

where $a(i)$ is the *tightness*, that is, the average distance from subject i to all other subjects in the same cluster, and $b(i)$ is the *degree of separation*, that is, the shortest average distance from subject i to any other cluster. The overall clustering quality is obtained by calculating the average Silhouette score across all data points. The range of a silhouette coefficient is between -1 and $+1$.

The Calinski-Harabasz index (CHI) [24] assesses how well-defined and distinct the clusters are, according to their *intra-cluster* and *inter-cluster variability*, and is computed as:

$$CHI = \frac{BCSS/(k-1)}{WCSS/(n-k)}, \tag{4}$$

where WCSS is the sum of squared distances from each point to its cluster mean, BCSS computes the sum of squared distances between each cluster mean and the overall mean, weighted by the number of data points in each cluster, n is the total number of data points, and k is the number of clusters. High CHI values indicate that the clusters are well-separated and compact. Low CHI values indicate that the clusters are less distinct and may overlap. A lower value suggests poor clustering performance.

Finally, the Davies-Bouldin Index [25] is a metric based on the average similarity ratio of each cluster with its most similar (i.e., closest) cluster, computed as follows:

$$DBI = \frac{1}{k} \sum_{i=1}^{k} \max_{j \neq i} R_{ij}, \tag{5}$$

where R_{ij} is the ratio of the inter-cluster distance to the intra-cluster distance, and k is the number of clusters.

While in the classical approach SS, CHI, and DBI are evaluated according to the distance between data points, here, since we are working with longitudinal data, these indexes are evaluated with respect to the distance between a longitudinal patient profile and a set of multiple profiles.

Furthermore, to evaluate the relevance of the obtained subjects' stratification, the Adjusted Rand Index metric [26] is applied as an external assessment metric considering the 28-day death feature as the target variable. This feature is considered as a traditional endpoint in several medical studies [22, 36, 37]. The Adjusted Rand Index (ARI) evaluates the similarity between two clustering outcomes with values in the $[-1; +1]$ range,

where +1 means perfect similarity, −1 means opposite similarity, and 0 means similarity no better than random chance, computed as follows:

$$ARI = \frac{RI - E}{1 - E} \tag{6}$$

where RI is the Rand Index and E is the expected value for the Rand Index [38].

Clinical variables identification and assessment

In this section, we report the results of the experimental analysis carried out to identify the clinical variables that better stratify subjects in terms of sepsis progression. In addition, we perform a comparison analysis to assess the obtained results with those obtained applying other longitudinal clustering methods.

Longitudinal clustering results

The multivariate longitudinal clustering method has been applied to the cohort of individuals defined in Dataset, to identify clinical sub-types in septic patients with different outcomes during ICU stays. Several tests were run by varying the number of clusters and the number and combinations of features.

To evaluate the robustness of the clustering performance, each configuration was tested across ten repeated runs, each involving a different shuffle of the dataset to simulate natural variability in data ordering and composition. From these repeated tests, mean performance, standard deviation (SD), and 95% confidence intervals (CI) were calculated, along with observed minimum and maximum values for each metric. Tables 3, 4, 5 and 6 report these values for the Silhouette Score (SS), Calinski-Harabasz Index (CHI), Davies-Bouldin Index (DBI), and Adjusted Rand Index (ARI) for those feature combinations that obtained the highest values. As we can see, the best result is obtained by stratifying into two clusters and considering two features, the Thrombin-Antithrombin complex and the Prothrombin Time–International Normalized Ratio.

The Silhouette Score (SS) (see Table 3) was particularly high, with a mean of 0.848, indicating excellent separation between clusters and tight cohesion within them. The associated standard deviation (±0.004) reflects low variability in performance across repeated runs with shuffled datasets, suggesting consistent cluster composition. The 95% confidence interval (CI) of [0.844, 0.854], combined with the observed maximum value of 0.854, confirms that the clustering quality is not only high but also stable across repeated experiments under slightly varied input conditions.

The Davies-Bouldin Index (DBI) (see Table 4), which penalizes poor separation and dispersion (lower values indicate better clustering), yielded a mean of 0.561 ± 0.024, with a 95% CI of [0.526–0.586]. These values support the conclusion that the clusters were compact and well-separated, with limited variability in quality across the different dataset shuffles.

Similarly, the Calinski-Harabasz Index (CHI) (see Table 5), which evaluates the ratio of between-cluster to within-cluster dispersion (higher is better), further confirmed the robustness of the clustering. The mean CHI was 191.5 ± 30.2, with a 95% CI of

[147.9–223.7]. Although this metric showed more variability than SS and DBI, the high values indicate that the clusters were clearly defined and statistically meaningful.

Hence, reaching a Silhouette Score as high as 0.854, a Calinski-Harabasz Index of 223.7, and a Davies-Bouldin Index as low as 0.526 demonstrates that the clustering can achieve exceptionally well-defined patient groupings, reinforcing the robustness and clinical promise of using TAT and INR as stratifying biomarkers.

Beyond internal validation, the Adjusted Rand Index (ARI), an external validation metric, was computed to evaluate the prognostic significance of the unsupervised clusters identified through the proposed longitudinal analysis. For this analysis the 28-day mortality binary variable was considered as target variable allowing us to measure how well the unsupervised clusters align with actual patient survival.

Results shown in Table 6 indicate that our clustering approach demonstrates discriminative ability when applied to longitudinal trajectories of the Thrombin–Antithrombin complex and Prothrombin Time-International Normalized Ratio. Specifically, the ARI showed a mean of 0.483 ± 0.023, a 95% CI of [0.464–0.511], and a maximum observed value of 0.511. While the mean value reflects a moderate alignment between clustering and mortality outcome, the maximum ARI of 0.511 indicates that clustering can capture over half of the variance in mortality outcome in groups, despite being derived in an unsupervised manner. This suggests that the TAT and INR trajectories encode meaningful prognostic information and may help identify sepsis subgroups with distinct survival profiles, which supports the hypothesis that early coagulation dynamics captured by TAT and INR may reflect underlying pathophysiological differences relevant to mortality risk.

Table 3 Silhouette score (SS) for $k = 2, 3, 4$. SS interval: $[-1; +1]$, the higher the better. The Max value, the mean, the standard deviation (SD), and the 95% confidence interval (CI) obtained during the tests are reported. *** Best result

# Features	Features	Silhouette Score		
		Max	Mean±SD	95% CI
Number of clusters= 2				
2	TAT, INR***	0.854	0.848 ± 0.004	[0.844, 0.854]
3	TAT, INR, FDP	0.831	0.810 ± 0.022	[0.786, 0.831]
4	TAT, INR, FDP, PlC	0.724	0.515 ± 0.223	[0.214, 0.724]
5	TAT, INR, PlC, AlPC, ProteinC	0.734	0.557 ± 0.197	[0.315, 0.729]
6	TAT, INR, PlC, FDP, AlPC, ProteinC	0.700	0.481 ± 0.210	[0.261, 0.697]
Number of clusters= 3				
2	INR, FDP	0.842	0.821 ± 0.062	[0.681,0.842]
3	TAT, INR, FDP	0.797	0.646 ± 0.124	[0.508, 0.797]
4	TAT, INR, FDP, ProteinC	0.765	0.529 ± 0.192	[0.297,0.765]
5	TAT, INR, PlC, AlPC, FDP	0.685	0.408 ± 0.165	[0.259, 0.685]
6	TAT, INR, PlC, FDP, AlPC, ProteinC	0.291	0.273 ± 0.032	[0.200, 0.291]
Number of clusters= 4				
2	INR, FDP	0.833	0.678 ± 0.132	[0.525, 0.826]
3	TAT, INR, FDP	0.708	0.553 ± 0.055	[0.498, 0.676]
4	TAT, INR, FDP, ProteinC	0.758	0.418 ± 0.133	[0.278, 0.695]
5	INR, PlC, AlPC, FDP, ProteinC	0.354	0.318 ± 0.035	[0.254, 0.353]
6	TAT, INR, PlC, FDP, AlPC, ProteinC	0.304	0.289 ± 0.018	[0.250, 0.304]

Table 4 Davies-Bouldin Index for $k = 2, 3, 4$. DBI interval: $[0; +\infty)$, the lower the better. The Min value, the mean, the standard deviation (SD), and the 95% confidence interval (CI) value obtained during the tests are reported. *** Best result

# Features	Features	Davies-Bouldin Index		
		Min	Mean±SD	95% CI
Number of clusters= 2				
2	TAT, INR ***	0.526	0.561 ± 0.024	[0.526, 0.586]
3	TAT, INR, FDP	0.540	0.574 ± 0.035	[0.540, 0.614]
4	TAT, INR, FDP, PlC	0.634	1.173 ± 0.580	[0.634, 1.966]
5	TAT, INR, PlC, AIPC, ProteinC	0.493	1.043 ± 0.615	[0.506, 1.811]
6	TAT, INR, PlC, FDP, AIPC, ProteinC	0.578	1.266 ± 0.658	[0.578, 2.010]
Number of clusters= 3				
2	INR, FDP	0.792	0.817 ± 0.075	[0.792, 0.986]
3	TAT, INR, FDP	1.003	1.073 ± 0.112	[1.004, 1.327]
4	TAT, INR, FDP, ProteinC	1.112	1.372 ± 0.267	[1.124, 1.945]
5	TAT, INR, PlC, AIPC, FDP	1.146	1.405 ± 0.179	[1.146, 1.684]
6	TAT, INR, PlC, FDP, AIPC, ProteinC	1.328	1.631 ± 0.489	[1.328, 2.581]
Number of clusters= 4				
2	INR, FDP	0.964	1.091 ± 0.093	[0.972, 1.201]
3	TAT, INR, FDP	1.023	1.252 ± 0.193	[1.044, 1.668]
4	TAT, INR, FDP, ProteinC	1.016	1.506 ± 0.216	[1.066, 1.733]
5	INR, PlC, AIPC, FDP, ProteinC	0.929	1.468 ± 0.234	[1.021, 1.821]
6	TAT, INR, PlC, FDP, AIPC, ProteinC	1.034	1.356 ± 0.201	[1.075, 1.692]

Table 5 Calinski-Harabasz index for $k = 2, 3, 4$. CHI interval: $[0; +\infty)$, the higher the better. The Max value, the mean value, the standard deviation (SD), and the 95% confidence interval (CI) obtained during the tests are reported. *** Best result

# Features	Features	Calinski-Harabasz Index		
		Max	Mean±SD	95% CI
Number of clusters= 2				
2	TAT, INR***	223.699	191.482 ± 30.157	[147.927, 223.699]
3	TAT, INR, FDP	111.901	77.725 ± 25.591	[53.687, 111.901]
4	TAT, INR, FDP, PlC	96.416	64.220 ± 31.188	[18.928, 96.416]
5	TAT, INR, PlC, AIPC, ProteinC	133.185	89.776 ± 48.501	[30.398, 133.185]
6	TAT, INR, PlC, FDP, AIPC, ProteinC	157.448	63.080 ± 46.114	[25.191, 145.146]
Number of clusters= 3				
2	INR, FDP	217.089	204.088 ± 39.002	[116.334, 217.089]
3	TAT, INR, FDP	137.187	107.036 ± 27.150	[52.401, 137.187]
4	TAT, INR, FDP, ProteinC	143.180	104.647 ± 37.228	[33.953, 142.542]
5	TAT, INR, PlC, AIPC, FDP	78.268	59.214 ± 16.501	[28.854, 78.268]
6	TAT, INR, PlC, FDP, AIPC, ProteinC	75.381	58.156 ± 26.279	[15.110, 75.381]
number of clusters= 4				
2	INR, FDP	194.999	171.282 ± 21.513	[139.146, 194.883]
3	TAT, INR, FDP	138.773	111.858 ± 24.341	[67.861, 138.331]
4	TAT, INR, FDP, ProteinC	130.226	98.612 ± 15.154	[79.989, 127.513]
5	INR, PlC, AIPC, FDP, ProteinC	110.361	67.450 ± 21.936	[33.595, 104.467]
6	TAT, INR, PlC, FDP, AIPC, ProteinC	80.620	68.335 ± 9.396	[52.842, 80.572]

Temporal phenotyping of septic patients

According to the previous clustering results, the optimal clusterization is obtained with k=2 along with TAT and INR variables. Figure 3 illustrates the trajectories of the Thrombin-Antithrombin complex and Prothrombin Time–International Normalized Ratio over the seven-day ICU period, corresponding to the patient profiles categorized within two clusters (i.e., Cluster 0 and Cluster 1) determined by using the multivariate longitudinal cluster described in Multivariate longitudinal cluster analysis section. The observed trend indicates that ICU patients categorized within Cluster 0 exhibit, on average, elevated levels of Thrombin-Antithrombin Complex and Prothrombin Time–International Normalized Ratio upon admission compared to individuals in Cluster 1. The INR values remained elevated for several days during the patient's stay in the intensive care unit. Then, these values deteriorated to a state deemed irreversible.

Conversely, ICU patients in Cluster 1 exhibit a more stable trajectory characterized, on average, by lower values than individuals in Cluster 0.

To better characterize subjects assigned to the two clusters, static variables such as average age, gender, and some clinical variables, the source of sepsis, and the two primary patient outcomes of most relevant significance, are presented in Table 7.

In such a table, the 28-day death variable is also reported. The distinctions among the clusters are further elucidated through the application of statistical methods, which assess their levels of statistical significance. To examine continuous data in relation to the normality of distributions, either analysis of variance (ANOVA) or the Kruskal-Wallis test was utilized. The chi-square test was exploited to assess the differences in the frequencies of categorical data. The significance level was set at p-value < 0.01.

Table 6 Adjusted Rand Index for $k = 2, 3, 4$. ARI interval: $[-1; +1]$, the higher the better. The Max value, the mean, the standard deviation (SD), and the 95% confidence interval (CI) are reported as obtained during the tests. *** Best result

# Features	Features	Adjusted Rand Index		
		Max	Mean±SD	95% CI
Number of clusters= 2				
2	TAT, INR***	0.511	0.483 ± 0.023	[0.464, 0.511]
3	TAT, INR, FDP	0.511	0.238 ± 0.255	[−0.016, 0.511]
4	TAT, INR, FDP, PIC	0.511	0.277 ± 0.230	[−0.031, 0.511]
5	TAT, INR, PIC, AIPC, ProteinC	0.464	0.277 ± 0.203	[0.027, 0.464]
6	TAT, INR, PIC, FDP, AIPC, ProteinC	0.464	0.199 ± 0.213	[−0.014, 0.464]
Number of clusters= 3				
2	INR, FDP	0.470	0.450 ± 0.062	[0.311, 0.470]
3	TAT, INR, FDP	0.470	0.300 ± 0.139	[0.092, 0.470]
4	TAT, INR, FDP, ProteinC	0.47	0.295 ± 0.146	[0.095, 0.470]
5	TAT, INR, PIC, AIPC, FDP	0.425	0.203 ± 0.155	[−0.002, 0.425]
6	TAT, INR, PIC, FDP, AIPC, ProteinC	0.102	0.073 ± 0.042	[−0.005, 0.102]
Number of clusters= 4				
2	INR, FDP	0.469	0.311 ± 0.121	[0.173, 0.458]
3	TAT, INR, FDP	0.368	0.223 ± 0.052	[0.175, 0.340]
4	TAT, INR, FDP, ProteinC	0.469	0.215 ± 0.104	[0.086, 0.421]
5	INR, PIC, AIPC, FDP, ProteinC	0.124	0.076 ± 0.040	[−0.008, 0.120]
6	TAT, INR, PIC, FDP, AIPC, ProteinC	0.115	0.094 ± 0.012	[0.072, 0.112]

As we can see, there is no statistical difference in age and gender between Cluster 0 and Cluster 1. However, in Cluster 0 there are statistically more individuals with sepsis caused by pulmonary infections than other sources of sepsis. The significant differences between patients in Cluster 0 and patients in Cluster 1 are related to the severity of their conditions. Patients in Cluster 0 show the worst APACHE II score and SOFA score with respect to patients in Cluster 1 (p-value $= 4.6 \cdot 10^{-5}$ and p-value $= 1.9 \cdot 10^{-3}$). The most noteworthy finding was that no individuals classified within Cluster 0 survived beyond 28 days. The 28-day death measure indicates that patients in Cluster 1 have a better chance of survival than patients in Cluster 0 (p-value $= 1.3 \cdot 10^{-18}$). The duration between the patient's admission to the intensive care unit and the event's occurrence was remarkably short for patients in Cluster 0: 5 days on average. On the contrary, the event for patients in Cluster 1 occurred on average after 20 days of ICU admission.

Survival Analysis of Clustered Patient Trajectories

To further evaluate the prognostic significance of the unsupervised clusters identified through longitudinal analysis of hemostatic and thrombopoietin biomarkers (i.e., INR and TAT), we also performed a survival analysis using the Kaplan–Meier estimator. Specifically, we assessed the 28-day survival probability across the two clusters (Cluster 0 and Cluster 1) obtained with the optimal number of clusters. For each patient, the survival time was defined as the number of days from ICU admission until either in-hospital death (event) or censoring at 28 days (survival).

Survival curves for each cluster were estimated using the Kaplan–Meier method, and the log-rank test was employed to statistically compare survival distributions between the two clusters [39]. Survival estimation and hypothesis testing were implemented via the lifelines package [40], which provides robust tools for time-to-event analysis in a flexible and reproducible framework.

As illustrated in Fig. 4, the Kaplan–Meier curves reveal a clear difference in survival probabilities between the two clusters. Notably, patients in Cluster 0 exhibited a markedly lower survival probability across the 28-day period compared to Cluster 1. The log-rank test yielded a p-value < 0.01, indicating that the observed difference in survival between clusters is statistically significant. These findings suggest that the patients' subgroups identified by differences in coagulation biomarker dynamics present different mortality risks.

Comparison with different longitudinal approaches

To further validate our results using multivariate longitudinal k-means, we selected three complementary clustering approaches—`mclust`, `fdapace`, and `gbmt`—to represent distinct methodological paradigms widely applied in longitudinal healthcare research. Mainly, `mclust` [41] is a model-based clustering method based on parameterized finite Gaussian mixture models. Recently, it was applied to cluster patients in chronic pain rehabilitation, revealing subgroups with distinct service utilization patterns [42] and in cervical cancer to stratify tumor budding and identify prognostically relevant EMT-related genes [43]. It is defined as a gold standard in model-based clustering [44].

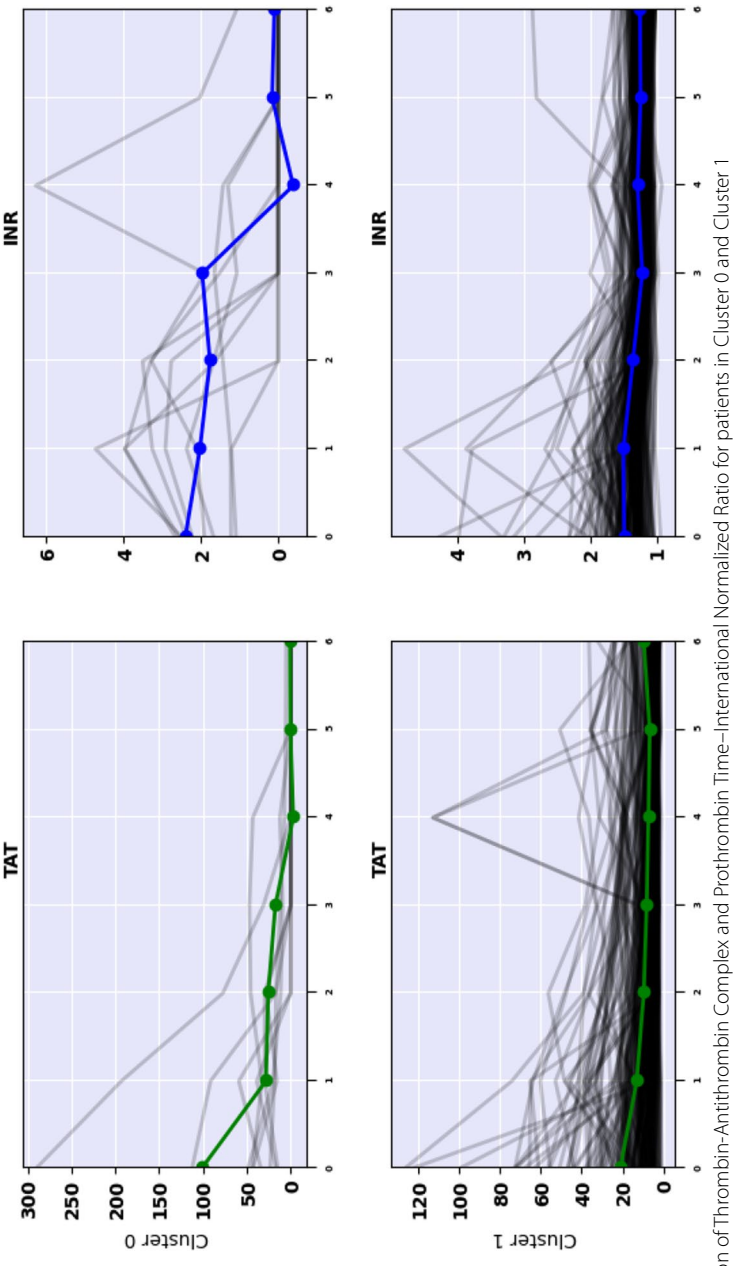

Fig. 3 Joint progression of Thrombin-Antithrombin Complex and Prothrombin Time–International Normalized Ratio for patients in Cluster 0 and Cluster 1

Table 7 Demographics and clinical characteristics of ICU patients. * One-way ANOVA, † Kruskal-Wallis test, ‡chi-square test

Variables	Cluster 0 ($n = 9$)	Cluster 1 ($n = 195$)	*p*-value
Age (years)			$3.1 \cdot 10^{-1}$†
mean ± SD	64.9 ± 12.9	68.5 ± 14.6	
[min, max]	[34, 81]	[19, 101]	
Gender			$1.6 \cdot 10^{-1}$‡
Male (%)	3 (33.3%)	114 (58.5%)	
Female (%)	6 (66.7%)	81 (41.5%)	
APACHE II score			$4.6 \cdot 10^{-5}$†
mean ± SD	39.3 ± 8.8	24.6 ± 7.5	
[min, max]	[30, 53]	[7, 42]	
SOFA score			$1.9 \cdot 10^{-3}$*
mean ± SD	11.7 ± 3.9	8 ± 3.3	
[min, max]	[4, 17]	[2, 19]	
ICU days			$3.9 \cdot 10^{-4}$†
mean ± SD	4.6 ± 4.8	10.9 ± 7.7	
[min, max]	[2, 17]	[2, 58]	
Source of sepsis			$4.6 \cdot 10^{-60}$‡
Pulmonary (%)	6 (66.7%)	44 (22.6%)	
Abdominal (%)	2 (22.2%)	99 (50.8%)	
Urinary tract (%)	0 (0%)	11 (5.6%)	
Soft tissue (%)	0 (0%)	26 (13.3%)	
Blood stream (%)	0 (0%)	2 (1%)	
other (%)	1 (11.1%)	13 (6.7%)	
Time to event			$8.6 \cdot 10^{-4}$†
mean ± SD	5.3 ± 6.5	19.9 ± 5.1	
[min, max]	[2, 22]	[11, 27]	
28-day death			$1.3 \cdot 10^{-18}$‡
survived (%)	0 (0%)	182 (93.3%)	
death (%)	9 (100%)	13 (6.7%)	

The `fdspace` method (Functional Data Analysis and Empirical Dynamics method) [45] applies functional principal component analysis to sparse and irregular longitudinal data. Among other applications, it has recently demonstrated its utility in obesity research. It has been used to analyze dynamic physical activity and metabolic biomarker trajectories, including modeling glycemic responses to prenatal exercise in obese pregnant women [46] and identifying circulating biomarkers mediating the link between circadian activity patterns and obesity risk in population cohorts [47]. It is also highlighting its relevance in clustering longitudinal viral load patterns for monitoring the HIV status by characterizing each cluster by demographics, comorbidities, and social behaviors [48]. Finally, `gbmt` (Group-Based Multivariate Trajectory Modeling) [49] models individual-level multivariate trajectories through latent class polynomial functions. It is commonly used to uncover patterns of symptom burden or behavioral change

in medical studies [50]. More recently, it has been applied for characterizing long-term medication adherence in patients with chronic illness [51].

These three methods collectively span probabilistic (`mclust`), functional (`fdapace`), and parametric trajectory-based (`gbmt`) approaches. Their demonstrated utility in healthcare research involving multivariate longitudinal data supports their inclusion as methodologically diverse and conceptually complementary benchmarks for assessing the performance of our multivariate longitudinal k-means approach.

In addition to methodological diversity, these methods were chosen for practical feasibility. All are supported by well-maintained and widely used R packages ensuring transparency, replicability, and computational accessibility.

Finally, by benchmarking our approach against well-established methods with demonstrated utility across diverse healthcare applications, we aim to determine whether it can yield clinically meaningful and interpretable groupings consistent with current methodological standards.

Comparison settings and results

To ensure a fair and consistent comparison across methods, we assessed the quality of each resulting clustering solution using ARI, which, as said before, quantifies the agreement between the clustering assignments and a predefined target variable. In this context, the 28-day mortality outcome was used, allowing us to evaluate how well each feature set and clustering method stratifies patients in a way that aligns with clinically relevant survival differences. To perform the comparison, we fixed the number of clusters to $k = 2$ according to the binary target variable. Thus, we evaluate and compare different clustering configurations by systematically varying the input feature combinations. The top five variable subsets identified by each method, along with the corresponding ARI, are listed in the corresponding rows of Table 8. Comparison of the Adjusted Rand Index across methods reveals that the stratification produced by our algorithm most effectively aligns with the 28-day mortality outcome, indicating superior predictive relevance.

The ARI values observed among the top five configurations support the robustness of our approach. Notably, the highest ARI score ($ARI = +0.51$) was achieved using only the Thrombin–Antithrombin complex (TAT) and Prothrombin Time–International Normalized Ratio (INR), demonstrating that even a minimal feature set can yield clinically meaningful stratification with our method.

Our algorithm's second-best performance ($ARI = +0.41$) was obtained using a slightly expanded feature set that included TAT, INR, platelet count, and FDP. In comparison, the `mclust` method [41] reached its best performance ($ARI = +0.49$) using Protein C, INR, and platelet count. Conversely, the `fdapace` and `gbmt` methods consistently underperformed across all tested feature combinations, indicating limited effectiveness in capturing survival-relevant stratification in this context.

It is also noteworthy that INR consistently emerges as one of the most predictive variable across all clustering algorithms, regardless of their overall performance. This recurrent selection—observed even in methods such as `fdapace` and `gbmt`, which otherwise yielded lower ARI scores—underscores the robustness and clinical

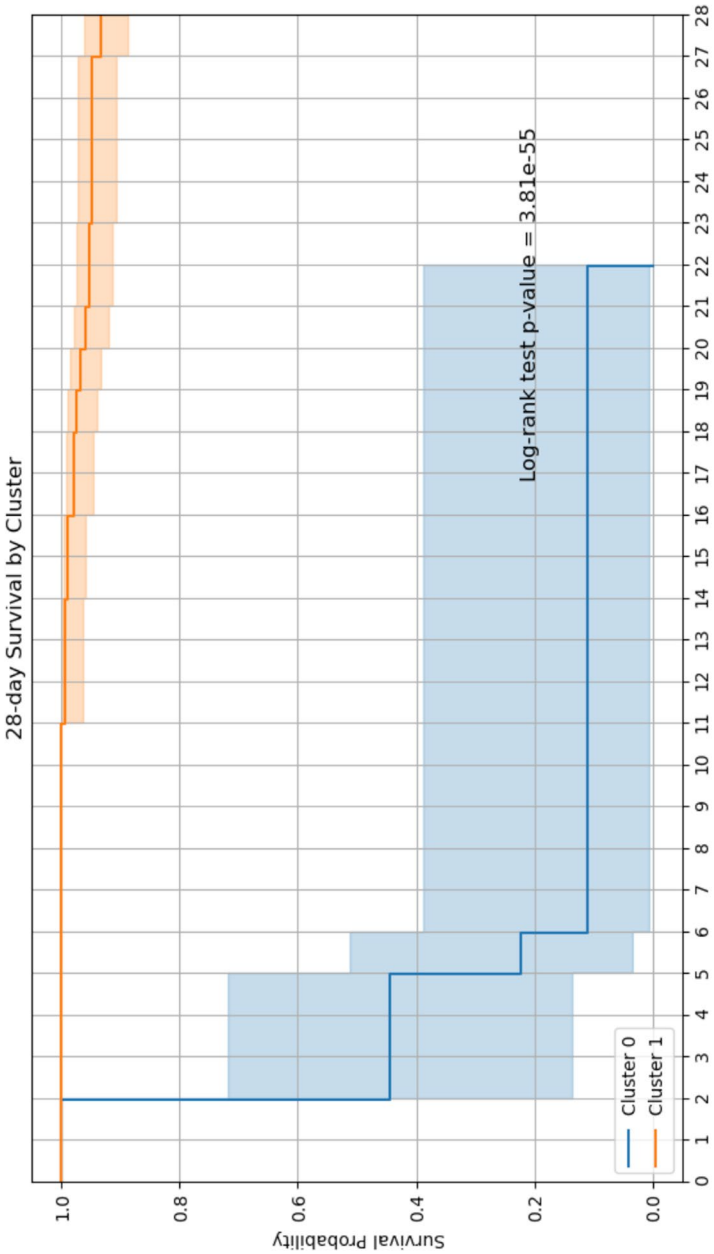

Fig. 4 Kaplan–Meier survival curves comparing 28-day mortality between patient clusters identified by longitudinal hemostatic and thrombopoietic biomarkers. Cluster 0 shows significantly lower survival probability compared to Cluster 1 (log-rank test, p = 3.81e-55). Time is expressed in days from ICU admission, with a step size of 1 day on the x-axis

relevance of INR as a key discriminator in stratifying septic patients based on 28-day mortality.

Discussion

This study focused on leveraging temporal clinical markers for stratifying septic patients according to their health status progression. We analyzed the longitudinal clinical variable measurements using an unsupervised clustering approach, which was opportunely developed for studying the temporal dynamics of patient health status changes. The adopted unsupervised approach detected two clusters of septic patients. TAT and INR, indicating a hypercoagulable and prothrombotic state, were identified as the two clinical variables most correlated to the different prognoses of ICU patients.

Sepsis is already known to be associated with Thrombin–Antithrombin complex [53–55] and with Prothrombin Time [56–58] in the scientific literature.

Our findings reveal that the cluster of patients with elevated levels of TAT and INR over time exhibits a significantly higher incidence of 28-day mortality compared to the other. These results confirm that the Thrombin-Antithrombin Complex and the Prothrombin Time–International Normalized Ratio are two biomarkers crucial in assessing and predicting sepsis progression. These markers help monitor coagulation and liver function, which are commonly affected in sepsis. Indeed, the Thrombin-Antithrombin Complex is a biomarker that indicates activation of coagulation. Thrombin is a key enzyme in blood clotting, while Antithrombin inhibits thrombin to prevent excessive clotting. When thrombin is activated during coagulation, it binds to antithrombin, forming the TAT complex. Elevated TAT levels in sepsis indicate hypercoagulation and may suggest the onset of disseminated intravascular coagulation (DIC) [59], a serious complication of sepsis in which widespread clotting and bleeding occur simultaneously. As the medical literature reveals, high TAT levels are often correlated with poor outcomes in septic patients, including a higher risk of organ dysfunction and mortality. Moreover, TAT elevation is associated with increased thrombotic complications and coagulopathy, which are risk factors for multi-organ failure.

On the contrary, INR measures blood coagulation calculated from the prothrombin time (PT), which measures how long it takes to clot [28]. It is often used to assess the coagulation pathway and monitor patients on anticoagulant therapy. Sepsis affects coagulation by promoting excessive clot formation and reducing the blood's ability to clot when needed. This leads to an imbalance between procoagulant and anticoagulant factors. Elevated INR indicates impaired coagulation and suggests a deficiency in clotting factors. Higher INR values are associated with increased bleeding risk and are often a sign of liver dysfunction in septic patients. Elevated INR is a predictor of worse outcomes in sepsis, especially when it is accompanied by organ failure or the development of DIC. In sepsis, INR elevation may signal progression toward severe sepsis or septic shock, where coagulopathies and multi-organ failure are common.

Since TAT and INR are critical markers of sepsis coagulation dysfunction, their combined assessment can provide valuable insight into a patient's risk of progressing to more severe stages of sepsis, such as severe sepsis, septic shock, or multi-organ failure.

Table 8 Relevance of the cluster stratifications with respect to the 28-day death target variable. Results obtained by applying our algorithm and other methods on a subset of clinical variables. *** Top performing Adjusted Rand Index. ARI: adjusted Rand index (interval: $[-1, +1]$, the higher the better)[52]

Ranking	Our algorithm on selected features:	ARI	ARI_{Mean}	ARI_{SD}	$ARI_{95\%_{low}}$	$ARI_{95\%_{up}}$
1	TAT, INR	***0.511	0.483	0.023	0.464	0.511
2	TAT, INR, platelet count, FDP	0.511	0.277	0.23	−0.031	0.511
3	TAT, INR, ProteinC	0.511	0.238	0.255	−0.016	0.511
4	TAT, INR, AIPC, FDP, ProteinC	0.464	0.277	0.203	0.027	0.464
5	TAT, INR, platelet count, AIPC, FDP, ProteinC	0.464	0.199	0.203	−0.014	0.464
`mclust` on selected features:						
1	ProteinC, INR, platelet count	0.486	0.2	0.14	−0.018	0.486
2	ProteinC, INR	0.447	0.172	0.131	−0.038	0.447
3	ProteinC, INR, AIPC, platelet count	0.413	0.219	0.123	0.033	0.413
4	ProteinC, INR, FDP, AIPC, platelet count	0.293	0.194	0.067	0.115	0.293
5	ProteinC, INR, TAT, FDP, AIPC, platelet count	0.163	0.167	0	0.163	0.163
`fdapace` on selected features:						
1	AIPC, FDP, TAT, INR	0.202	0.063	0.068	−0.015	0.202
2	FDP, TAT, ProteinC	0.165	0.025	0.068	−0.056	0.165
3	TAT, INR	0.153	0.002	0.059	−0.072	0.153
4	AIPC, FDP, TAT, INR, ProteinC	0.151	0.046	0.049	0.003	0.151
5	AIPC, FDP, TAT, INR, ProteinC, platelet count	0.027	0.027	0	0.027	0.027
`gbmt` on selected features:						
1	AIPC, INR	0.118	0.021	0.049	−0.023	0.118
2	AIPC, TAT, INR, ProteinC	0.128	0.007	0.040	−0.051	0.128
3	AIPC, TAT, INR	0.113	0.014	0.036	−0.025	0.113
4	AIPC, FDP, INR, ProteinC, platelet count	0.023	−0.003	0.014	−0.022	0.023
5	AIPC, FDP, TAT, INR, ProteinC, platelet count	−0.003	−0.003	0	−0.003	−0.003

As confirmed by our findings, a high TAT (indicating a pro-thrombotic state) and elevated INR (indicating impaired clotting) that have been present for days can indicate that a septic patient is at a critical point in disease progression. As such, patients belonging to this cluster might represent a distinct pathophysiological subgroup with a more adverse clinical outcome.

While our current analysis is exploratory and unsupervised, the identification of such a cluster of patients may suggest to clinicians to consider closer monitoring, more aggressive supportive care, or even early anticoagulant-based interventions for patients with the same clinical profile. Although our study was not designed to directly propose treatment modifications, the observed stratification may drive future risk-adapted interventions, such as intensified hemostatic monitoring or early consideration of adjunctive therapies targeting coagulation pathways. This approach may be especially relevant in the context of precision medicine in critical care, where tailoring interventions to biological phenotypes holds increasing promise.

Importantly, our clustering approach appears to capture prognostically relevant information that is not fully accounted for by standard severity scores. Specifically, both SOFA and APACHE II scores are well-established tools for assessing overall disease

severity and organ dysfunction, but they differ in purpose, composition, and computational complexity.

The SOFA (Sequential Organ Failure Assessment) score is designed to quantify the degree of organ dysfunction across six systems (respiratory, cardiovascular, hepatic, coagulation, renal, and neurological), requiring daily measurements of multiple physiological and laboratory parameters.

In contrast, the APACHE II (Acute Physiology and Chronic Health Evaluation II) score was explicitly developed as a mortality prediction model, incorporating 12 acute physiological variables (AaDO2 or PaO2, body temperature, arterial pressure blood pH, heart rate, respiratory rate, serum sodium, serum potassium, creatinine, hematocrit, white blood cell count, Glasgow Coma Scale), age and chronic health conditions (e.g., liver cirrhosis, portal hypertension, heart failure, severe respiratory disease, dialysis-dependent).

In our study, we observed that the clustering based on TAT and INR trajectories stratified mortality risk independently of baseline SOFA and APACHE II scores, suggesting that it provides complementary prognostic insight. In addition, these biomarkers, that directly implicated in sepsis-associated coagulopathy, are not explicitly quantified in SOFA or APACHE II.

Furthermore, the approach demonstrates prognostic utility in terms of both incidence rates and survival analysis. This is evident from the significant association with 28-day mortality, which is commonly adopted as a clinical endpoint for quantifying patient survival from sepsis [22, 36, 37].

It is particularly noteworthy that the Thrombin–Antithrombin complex (TAT) and the Prothrombin Time (INR) consistently emerge as among the most predictive features, not only when applying our approach, but also across different longitudinal clustering algorithms. These results suggest that INR and TAT not only play a significant role in distinguishing between patient subgroups over time but also remain robustly relevant regardless of the specific clustering method applied.

Finally, as it is shown in Fig. 3, the dispersion of individual trajectories around the cluster centroid indicates a level of heterogeneity that the current clustering solution may not fully capture. This deserves a further analysis of this cluster by exploiting different techniques such as hierarchical and fuzzy-based clustering approaches that could reveal more nuanced understanding of disease dynamics providing a finer patient stratification.

Conclusions

The study proposed in this paper leverages the potential of our longitudinal clustering approach to discover different profiles of septic patients by including the temporal aspect of clinical variables, delineating patients' health conditions. The obtained stratification proved to be consistent with the septic patient's survival prediction, as confirmed when adopting the 28-day death target variable.

The clinical variables most contributing to this stratification are the Thrombin–Antithrombin complex and Prothrombin Time–International Normalized Ratio, as the most critical ones mainly correlated with adverse outcomes. This stratification has the potential to facilitate the development of tailored therapeutic strategies, such as implementing more timely care protocols designed explicitly for high-risk cohorts of septic patients.

Moreover, considering the high-dimensional characteristics of electronic health records, particularly in the context of sepsis, which encompasses a diverse array of vital signs (such as heart rate, blood pressure, body temperature, and oxygen saturation), laboratory results (including white blood cell count, C-protein levels, platelet count, and Thrombin-Antithrombin complex, among others), as well as patient demographics (including age, sex, and underlying health conditions), the application of our multivariate longitudinal clustering proves advantageous in discerning patient subgroups that may exhibit distinct patterns of sepsis progression.

The robustness of the proposed approach was assessed by carrying out the same experiments with different longitudinal clustering algorithms. An external assessment metric was adopted to ensure an objective comparison. Notably, our algorithm consistently outperformed all other algorithms, highlighting its reliability and effectiveness in predicting survival outcomes.

Nevertheless, all considered algorithms revealed that the INR factor is among the most important features to be carefully monitored for the sepsis progression.

However, this study's major limitation is that its analytical framework relied exclusively on data derived from a single dataset. An analysis encompassing a larger dataset may enhance the validity of the generalizations derived from our findings. In addition, the reported results have to be supported by clinicians to assess their clinical validity. For future work, we will extend our approach to include probabilistic and fuzzy methods. This could be more helpful when the boundaries between clusters are not well-defined, as may occur in medical settings.

Abbreviations

AIPC	Absolute immature platelet counts
APACHE II	Acute Physiology and Chronic Health Evaluation II
CHI	Calinski-Harabasz index
CHF	Congestive heart failure
CKD	Chronic kidney disease
COPD	Chronic obstructive pulmonary disease
CVD	Cardiovascular disease
DBI	Davies-Bouldin index
DIC	Disseminated intravascular coagulation
DTW	Dynamic Time Warping
EHRs	Electronic health records
EM	Expectation-maximization
FAIR	Findability, accessibility, interoperability, and reusability
FDP	Fibrin degradation product
ICU	Intensive care unit
IHD	Ischemic heart disease
INR	Prothrombin time–international normalized ratio
LSTM	Long short-term memory
ML	Machine learning
S	Silhouette
SIRS	Systemic Inflammatory Response Syndrome
SOFA	Sequential organ failure assessment
TAT	Thrombin–Antithrombin complex
UTIs	Urinary tract infections

Software code availability

The code that implements Algorithm 1 is publicly available on GitHub (https://github.com/PatriziaRibino/LongitProgression) under GNU-GPL.

Authors'contributions

Contributions All authors conceived the proposed study and reviewed the manuscript. P.R. designed the algorithm, preprocessed data, performed the tests, supervised the study, and contributed to the article's writing. M.M. supervised the study, preprocessed data, and contributed to the article's writing. C.D.N. supervised the study, preprocessed data,

and contributed to the article's writing. G.P. worked on the algorithm, preprocessed data, and contributed to the article's writing. D.C. collected the dataset, supervised the study, and contributed to the article's writing. F.C. supervised the study, performed the literature review, and contributed to the article's writing.

Funding
This study was developed within the project funded by Next Generation EU – "Age-It – Ageing Well in an Ageing Society" project (PE0000015), National Recovery and Resilience Plan (NRRP) – PE8 – Mission 4, C2, Intervention 1.3, CUP B83C22004880006.

Data availability
The dataset used in this paper is publicly available under the Creative Commons Attribution 4.0 International Deed (CC BY 4.0) license on Figshare at the following URL: https://figshare.com/articles/dataset/Time_course_of_immature_platelet_count_and_its_relation_to_thrombocytopenia_and_mortality_in_patients_with_sepsis/5837823?file=10343616.

Declarations

Ethics approval and consent to participate
The authorization for collecting the data from patients and releasing them publicly was obtained by the original dataset curators [22].

Competing interests
The authors declare no competing interests.

Received: 13 June 2025 Accepted: 26 August 2025
Published online: 26 September 2025

References

1. Hotchkiss RS, Moldawer LL, Opal SM, Reinhart K, Turnbull IR, Vincent JL. Sepsis and septic shock. Nat Rev Dis Prim. 2016;2(1):1–21. https://doi.org/10.1056/nejmc1312359.
2. Davies M, Hagen PO. Systemic inflammatory response syndrome. Br J Surg. 1997;84(7):920–35. https://doi.org/10.1002/bjs.1800840707.
3. Tsantes AG, Parastatidou S, Tsantes EA, Bonova E, Tsante KA, Mantzios PG, et al. Sepsis-induced coagulopathy: an update on pathophysiology, biomarkers, and current guidelines. Life (Basel). 2023;13(2):350. https://doi.org/10.3390/life13020350.
4. Sun GD, Zhang Y, Mo SS, Zhao MY. Multiple organ dysfunction syndrome caused by sepsis: risk factor analysis. Int J Gen Med. 2021. https://doi.org/10.2147/ijgm.s328419.
5. Klein Klouwenberg PM, Spitoni C, van der Poll T, Bonten MJ, Cremer OL. Predicting the clinical trajectory in critically ill patients with sepsis: a cohort study. Crit Care. 2019;23:1–9. https://doi.org/10.1186/s13054-019-2687-z.
6. Zhao X, Shen W, Wang G. Early prediction of sepsis based on machine learning algorithm. Comput Intell Neurosci. 2021;2021(1):6522633. https://doi.org/10.1155/2021/6522633.
7. Fagerström J, Bång M, Wilhelms D, Chew M. LiSep LSTM: a machine learning algorithm for early detection of septic shock. Sci Rep. 2019;9:15132. https://doi.org/10.1038/s41598-019-51219-4.
8. M Amaral D, Soares DF, Gromicho M, de Carvalho M, Madeira SC, Tomás P, et al. Temporal stratification of amyotrophic lateral sclerosis patients using disease progression patterns. Nat Commun. 2024;15(1):5717.
9. Ramamoorthy D, Severson K, Ghosh S, Sachs K, Baxi EG, Coyne AN, et al. Identifying patterns in amyotrophic lateral sclerosis progression from sparse longitudinal data. Nat Comput Sci. 2022;2(9):605–16. https://doi.org/10.1038/s43588-022-00299-w.
10. Ribino P, Di Napoli C, Paragliola G, Chicco D, Gasparini F. Multivariate longitudinal clustering reveals neuropsychological factors as dementia predictors in an Alzheimer's disease progression study. BioData Min. 2025. https://doi.org/10.1186/s13040-025-00441-0.
11. Balk R, Esper AM, Martin GS, Miller RR, Lopansri BK, Burke JP, et al. Rapid and robust identification of sepsis using SeptiCyte RAPID in a heterogenous patient population. J Clin Med. 2024;13(20):6044. https://doi.org/10.3390/jcm13206044.
12. Zhang T, Wang S, Hua D, Shi X, Deng H, Jin S, et al. Identification of ZIP8-induced ferroptosis as a major type of cell death in monocytes under sepsis conditions. Redox Biol. 2024;69:102985. https://doi.org/10.1016/j.redox.2023.102985.
13. Knobloch CJ, Bourbia A. AI-driven sepsis mortality analysis: identifying phenotypic clusters using unsupervised machine learning. Chest. 2024;166(4):a6416. https://doi.org/10.1016/j.chest.2024.07.046.
14. Liu R, Greenstein JL, Fackler JC, Bembea MM, Winslow RL. Spectral clustering of risk score trajectories stratifies sepsis patients by clinical outcome and interventions received. eLife. 2020;9:e58142. https://doi.org/10.7554/elife.58142.
15. Papin G, Bailly S, Dupuis C, Ruckly S, Gainnier M, Argaud L, et al. Clinical and biological clusters of sepsis patients using hierarchical clustering. PLoS One. 2021;16(8):e0252793. https://doi.org/10.1371/journal.pone.0252793.
16. Kim M, Kym D, Hur J, Park J, Yoon J, Cho YS, et al. Tracking longitudinal biomarkers in burn patients with sepsis and acute kidney injury: an unsupervised clustering approach. Eur J Med Res. 2023;28(1):295. https://doi.org/10.1186/s40001-023-01268-3.

17. Yoon J, Kym D, Hur J, Cho YS, Chun W, Yoon D. Longitudinal profile of routine biomarkers for mortality prediction using unsupervised clustering algorithm in severely burned patients: a retrospective cohort study with prospectively collected data. Ann Surg Treat Res. 2023;104(2):126–35. https://doi.org/10.4174/astr.2023.104.2.126.
18. Genolini C, Falissard B. KmL: a package to cluster longitudinal data. Comput Methods Prog Biomed. 2011;104(3):e112–21. https://doi.org/10.1016/j.cmpb.2011.05.008.
19. Chicco D, Oneto L. Data analytics and clinical feature ranking of medical records of patients with sepsis. BioData Min. 2021;14(1):12. https://doi.org/10.1186/s13040-021-00235-0.
20. Mollura M, Chicco D, Paglialonga A, Barbieri R. Identifying prognostic factors for survival in intensive care unit patients with SIRS or sepsis by machine learning analysis on electronic health records. PLoS Digit Health. 2024;3(3):e0000459. https://doi.org/10.1371/journal.pdig.0000459.
21. Chicco D, Jurman G. Survival prediction of patients with sepsis from age, sex, and septic episode number alone. Sci Rep. 2020. https://doi.org/10.1038/s41598-020-73558-3.
22. Koyama K, Katayama S, Muronoi T, Tonai K, Goto Y, Koinuma T, et al. Time course of immature platelet count and its relation to thrombocytopenia and mortality in patients with sepsis. PLoS One. 2018;13(1):e0192064. https://doi.org/10.1371/journal.pone.0192064.
23. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7.
24. Calinski T, Harabasz J. A dendrite method for cluster analysis. Commun Stat-Theory Methods. 1974;3(1):1–27. https://doi.org/10.1080/03610927408827101.
25. Davies DL, Bouldin DW. A cluster separation measure. IEEE Trans Pattern Anal Mach Intel. 1979;PAMI-1(2):224–227. https://doi.org/10.1109/tpami.1979.4766909.
26. Zhang S, Wong HS, Shen Y. Generalized adjusted rand indices for cluster ensembles. Pattern Recognit. 2012;45(6):2214–26. https://doi.org/10.1016/j.patcog.2011.11.017.
27. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data. 2016. https://doi.org/10.1038/sdata.2016.18.
28. Doargaleh A, Favaloro E, Bahraini M, Rad F. Standardization of prothrombin time/international normalized ratio (PT/INR). Int J Lab Hematol. 2020;43(1). https://doi.org/10.1111/ijlh.13349.
29. Dahlbäck B, Villoutreix BO. The anticoagulant protein C pathway. FEBS Lett. 2005;579:3310–6. https://doi.org/10.1016/j.febslet.2005.03.001.
30. Likas A, Vlassis N, Verbeek JJ. The global k-means clustering algorithm. Pattern Recognit. 2003;36(2):451–61.
31. McLachlan G, Krishnan T. The EM algorithm and extensions. Wiley; 1997.
32. Berndt DJ, Clifford J. Using dynamic time warping to find patterns in time series. In: Proceedings of the 3rd international conference on knowledge discovery and data mining. 1994. pp. 359–370.
33. Sakoe H, Chiba S. Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans Acoust Speech Signal Process. 1978;26(1):43–9. https://doi.org/10.1109/tassp.1978.1163055.
34. Cuturi M, Blondel M. Soft-DTW: a differentiable loss function for time-series. In: Proceedings of the 34th International Conference on Machine Learning, PMLR 70, 2017. Sydney; 2017. https://proceedings.mlr.press/v70/cuturi17a/cuturi17a.pdf.
35. Tavenard R. Soft-DTW: A Differentiable Loss Function for Time-Series. 2021. https://rtavenar.github.io/blog/softdtw.html. Accessed 1 July 2025.
36. Karakike E, Kyriazopoulou E, Tsangaris I, Routsi C, Vincent JL, Giamarellos-Bourboulis EJ. The early change of SOFA score as a prognostic marker of 28-day sepsis mortality: analysis through a derivation and a validation cohort. Crit Care. 2019;23(1):387.
37. Patel JM, Snaith C, Thickett DR, Linhartova L, Melody T, Hawkey P, et al. Randomized double-blind placebo-controlled trial of 40 mg/day of atorvastatin in reducing the severity of sepsis in ward patients (ASEPSIS trial). Crit Care. 2012;16(6):R231.
38. Rand WM. Objective criteria for the evaluation of clustering methods. J Am Stat Assoc. 1971;66(336):846–50.
39. Kleinbaum DG. In: Kaplan-Meier Survival Curves and the Log-Rank Test. New York: Springer New York; 1996. pp. 45–82. https://doi.org/10.1007/978-1-4757-2555-1_2.
40. Davidson-Pilon C. Lifelines: survival analysis in Python. J Open Source Softw. 2019;4(40):1317. https://doi.org/10.21105/joss.01317.
41. Fraley C, Raftery AE, Scrucca L, Murphy TB, Fop M. mclust: Gaussian mixture modelling for model-based clustering, classification, and density estimation. 2025. https://doi.org/10.32614/cran.package.mclust. URL visited on 22nd April.
42. Strigo IA, Simmons AN, Giebler J, Schilling JM, Moeller-Bertram T. Unsupervised learning for prognostic validity in patients with chronic pain in transdisciplinary pain care. Sci Rep. 2023;13(1):7581.
43. Lee OE, Le TM, Chong GO, Cho JJ, Park NJY. The mclust analysis of tumor budding unveils the role of the collagen family in cervical cancer progression. Life (Basel). 2024;14(8):1004.
44. Corneli M, Erosheva E, Qian X, Lorenzi M, Alzheimer's Disease Neuroimaging Initiative. A bayesian approach for clustering and exact finite-sample model selection in longitudinal data mixtures. Comput Stat. 2025;40(1):509–45.
45. Zhou Y. fdapace: functional data analysis and empirical dynamics. 2025. https://doi.org/10.32614/cran.package.fdapace. URL visited on 22nd April.
46. Jansen LT, Stewart S, Cheak L, Jeffrey P, Andres A. Prenatal supervised physical activity maintains glycemic patterns in women with obesity during pregnancy. Curr Dev Nutr. 2024. https://doi.org/10.1016/j.cdnut.2024.103071.
47. Zhang Y, Xie L, Liu Y, Xing J, Xu J, Xu XS, et al. Identifying key circulating biomarkers that mediate the association between dynamic physical activity patterns and obesity risk. Discover Medicine. 2025;2(1):1–15.
48. Ma Y, Yang X, Xiao J, Li X, Olatosi B, Zhang J. Longitudinal Viral Load Clustering for People With HIV Using Functional Principal Component Analysis. AIDS Res Treat. 2025;2025(1):5890464.
49. Magrini A. gbmt: group-based multivariate trajectory modeling. 2025. https://doi.org/10.32614/cran.package.gbmt. URL visited on 22nd April.

50. Shi Q, Mendoza TR, Gunn GB, Wang XS, Rosenthal DI, Cleeland CS. Using group-based trajectory modeling to examine heterogeneity of symptom burden in patients with head and neck cancer undergoing aggressive non-surgical therapy. Qual Life Res. 2013;22:2331–9.
51. Cheruvu SS, Fatima B, Abughosh S. Group-based trajectory modeling to identify longitudinal patterns and predictors of adherence among older adults on concomitant triple therapy. J Manag Care Spec Pharm. 2025;31(4):396–405.
52. Hubert L, Arabie P. Comparing partitions. J Classif. 1985;2(1):193–218. https://doi.org/10.1007/bf01908075.
53. Koyama K, Madoiwa S, Nunomiya S, Koinuma T, Wada M, Sakata A, et al. Combination of thrombin-antithrombin complex, plasminogen activator inhibitor-1, and protein C activity for early identification of severe coagulopathy in initial phase of sepsis: a prospective observational study. Crit Care. 2014;18:1–11. https://doi.org/10.1186/cc13190.
54. Li J, Zhou J, Ren H, Teng T, Li B, Wang Y, et al. Clinical efficacy of soluble thrombomodulin, tissue plasminogen activator inhibitor complex, thrombin-antithrombin complex, *a*2-plasmininhibitor-plasmin complex in pediatric sepsis. Clin Appl Thromb Hemost. 2022;28:10760296221102928. https://doi.org/10.1177/10760296221102929.
55. Zhong L, Song X, WU J, Deng X, Lin Q, HE L, et al. Clinical significance of plasma thrombin-antithrombin complex levels in patients with sepsis-induced cardiomyopathy. Chin J Lab Med. 2022:1070–1075.
56. Schupp T, Weidner K, Rusnak J, Jawhar S, Forner J, Dulatahu F, et al. Diagnostic and prognostic significance of the prothrombin time/international normalized ratio in sepsis and septic shock. Clin Appl Thromb Hemost. 2022;28:10760296221137892. https://doi.org/10.1177/10760296221137893.
57. Kim SM, Kim SI, Yu G, Kim JS, Hong SI, Chae B, et al. Role of thromboelastography in the evaluation of septic shock patients with normal prothrombin time and activated partial thromboplastin time. Sci Rep. 2021;11(1):11833. https://doi.org/10.1038/s41598-021-91221-3.
58. Ishizuka M, Tago K, Kubota K. Impact of prothrombin time-International Normalized Ratio on outcome of patients with septic shock receiving polymyxin B cartridge hemoperfusion. Surgery. 2014;156(1):168–75. https://doi.org/10.1016/j.surg.2014.03.009.
59. Adelborg K, Larsen JL, Hvas AM. Disseminated intravascular coagulation: epidemiology, biomarkers, and management. Br J Haematol. 2021;192:803–18. https://doi.org/10.1111/bjh.17172.

Publisher's Note