SHOCK, Vol. 63, No. 4, pp. 573–578, 2025

INFORMING INTENSIVE CARE UNIT DIGITAL TWINS: DYNAMIC ASSESSMENT OF CARDIORESPIRATORY FAILURE TRAJECTORIES IN PATIENTS WITH SEPSIS

Grace Yao Hou,[1] Amos Lal,[2] Phillip J. Schulte,[3] Yue Dong,[4] Oguz Kilickaya,[2] Ognjen Gajic,[2] and Xiang Zhong[1]

[1]Department of Industrial and Systems Engineering, University of Florida, Gainesville, Florida; [2]Division of Pulmonary and Critical Care Medicine, Mayo Clinic College of Medicine and Science, Mayo Clinic, Rochester, Minnesota; [3]Division of Clinical Trials and Biostatistics, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota; [4]Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, Minnesota

Received 24 Sep 2024; first review completed 18 Oct 2024; accepted in final form 3 Dec 2024

ABSTRACT—Understanding clinical trajectories of sepsis patients is crucial for prognostication, resource planning, and to inform digital twin models of critical illness. This study aims to identify common clinical trajectories based on dynamic assessment of cardiorespiratory support using a validated electronic health record data that covers retrospective cohort of 19,177 patients with sepsis admitted to intensive care units (ICUs) of Mayo Clinic Hospitals over 8-year period. Patient trajectories were modeled from ICU admission up to 14 days using an unsupervised machine learning two-stage clustering method based on cardiorespiratory support in ICU and hospital discharge status. Of 19,177 patients, 42% were female with a median age of 65 (interquartile range [IQR], 55–76) years, The Acute Physiology, Age, and Chronic Health Evaluation III score of 70 (IQR, 56–87), hospital length of stay (LOS) of 7 (IQR, 4–12) days, and ICU LOS of 2 (IQR, 1–4) days. Four distinct trajectories were identified: fast recovery (27% with a mortality rate of 3.5% and median hospital LOS of 3 (IQR, 2–15) days), slow recovery (62% with a mortality rate of 3.6% and hospital LOS of 8 (IQR, 6–13) days), fast decline (4% with a mortality rate of 99.7% and hospital LOS of 1 (IQR, 0–1) day), and delayed decline (7% with a mortality rate of 97.9% and hospital LOS of 5 (IQR, 3–8) days). Distinct trajectories remained robust and were distinguished by Charlson Comorbidity Index, The Acute Physiology, Age, and Chronic Health Evaluation III scores, as well as day 1 and day 3 SOFA ($P < 0.001$ ANOVA). These findings provide a foundation for developing prediction models and digital twin decision support tools, improving both shared decision making and resource planning.

KEYWORDS—Sepsis; longitudinal clustering; clinical trajectory; cardiorespiratory failure

ABBREVIATIONS—IQR — interquartile range; LOS — length of stay; EHR — electronic health records; SOFA — Sequential Organ Failure Assessment; DTW — dynamic time warping; CCI — Charlson Comorbidity Index; SD — standard deviation

INTRODUCTION

Sepsis is a life-threatening condition caused by an extreme response to an infection, which can lead to tissue damage, organ failure, and death. In the United States, sepsis contributes to 35% of hospital deaths, with hospital costs for sepsis care exceeding $62 billion annually (1). The 30-day septic shock mortality rate is around 33.7% (2). The extent of organ dysfunction and requirements for organ support are pivotal in determining the severity and prognosis of sepsis. Specifically, cardiorespiratory failure is the most prevalent organ dysfunctions in sepsis and contributes significantly to the cascade of events leading to other organ failures and overall outcome: respiratory failure occurs in 82.6% of patients during their intensive care unit (ICU) stay, followed by cardiovascular failure in 45.5% of patients (3,4).

Given the clinical heterogeneity of sepsis, it is crucial to have a nuanced understanding of variability, including the degree of organ failure and need for organ support. Trajectory clustering, which groups patients based on the evolution of their clinical status over time, offers a powerful tool to dissect this variability. Although sepsis phenotypes have previously been discussed in the literature and have commonly been seen in clinical practice (5–7), our work focuses on presenting the trajectory of patients with sepsis akin to the phenotypes, which could provide the bedside clinicians with prognostic enrichment on patients' clinical progress and healthcare system needs. This is closely related to an emerging initiative to create digital twins of critically ill patients and the care delivery system (i.e., the patient-level digital twin and the ICU level digital twin environment) to enhance medical education, *in silico* research, and clinical decision support by using models informed by statistical and machine learning methods (8–11). The prototypical models of patients with sepsis in such digital twins considered patients as a homogeneous population. By clustering patient trajectories, subphenotypes of sepsis that share common progression patterns can be identified, and incorporating these phenotypes into the digital twins could enhance the fidelity of the models, facilitating the delivery of individualized therapeutic interventions. This approach can further facilitate goals of care discussions when communicating with caregivers, as specific trajectories may inform likely outcomes and treatment pathways. Additionally, trajectory clustering aids in resource management by predicting the expected course of care, allowing for more efficient allocation of healthcare resources and better planning for intensive care unit needs (11).

Address reprint requests to: Xiang Zhong, PhD, 303 Weil Hall, P.O. Box 116595, Gainesville, FL 32611-6595. E-mail: xiang.zhong@ise.ufl.edu

Financial support and conflict of interest disclosure: This study was supported by the National Science Foundation (Grant IIS-2123900 and IIS-2123848).

The authors report no conflicts of interest.

Funding: This study was supported by the National Science Foundation (Grant IIS-2123900 and IIS-2123848).

Supplemental digital content is available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.shockjournal.com).

DOI: 10.1097/SHK.0000000000002536

Many advancements in clustering techniques have enabled the identification of sepsis subclasses or subphenotypes (12). Subclass identification has been based mostly on static data, including clinical, genomic, or transcriptomic profiles. Genomic and transcriptomic-based clustering provides biological insights and robust risk stratification and reveals underlying mechanisms (13–15), while clinical data focuses on demographic variables and the most abnormal values of vital signs, laboratory results, and organ dysfunctions (5). Additionally, sepsis is a dynamic process, and static snapshots of clinical data capture its evolving nature with limited granularity and accuracy. Recent research has utilized longitudinal data leveraging electronic health records (EHR) to capture this dynamic nature by identifying trajectories based on vital signs (6,7,16–18) and Sequential Organ Failure Assessment (SOFA) scores (19). However, these studies have primarily focused on short time windows, ranging from 8 (18) h, 12 h (16), to 72 h (6,7,17,19) after admission, which ignored condition changes after that especially for patients with prolonged ICU stay. In addition, organ dysfunction was not fully explored in clustering studies for patients with sepsis. The existing study used the SOFA score to depict the level of organ dysfunction, which assumes equal importance for the six organ systems, and it is calculated at a low time resolution of 6 h (19).

To date, there is a lack of research depicting the comprehensive clinical trajectory of patients with sepsis from the time of admission until discharge or death with high time resolution, delineating the fluctuations in cardiorespiratory failure and required organ support during their ICU stay and hospital discharge status. Our study addressed this gap using granular EHR data to create clinical trajectories with a high time resolution of 1 h, allowing for close tracking of the clinical progression in patients with sepsis. Additionally, we chose a 14-day observation period, allowing us to capture a significant trajectory that could contribute to a loss of independence or poor quality of life (even after the survival from critical illness) as patients have protracted hospital stays. Our objective was to delineate clinical trajectories for patients with sepsis throughout the hospital stay and categorize them into clinically relevant clusters. Using two-stage clustering combining longitudinal K-means clustering and dynamic time warping (DTW), this study identified four major clinically relevant trajectories of patients with sepsis based on the cardiorespiratory failure and discharge status. Combined with prognostic enrichment, this work is valuable in clinical settings for shared decision making, resource utilization planning, and discussions regarding care goals.

MATERIALS AND METHODS

Study population

This study was approved by the institutional review board of the Mayo Clinic (18-000831), ensuring compliance with ethical standards and the protection of patient data. The study used EHR data from Mayo Clinic on 19,177 ICU patients with sepsis from 2011 to 2018. Patients with sepsis were identified if they met the criteria aligned with Sepsis-3 guidelines (20,21). The exclusion criteria for this study include 1) patients who did not provide research authorization required by Minnesota state, 2) below the age of 18, and 3) those with more than 24 h of pre-ICU care or was readmitted to ICU, as we want to focus on early sepsis cases. Our goal was to ensure that the cohort comprised patients who were admitted into ICU and confirmed to have sepsis in its early stages, without the influence of complications arising from an extended hospital stay prior to ICU admission. This approach allowed us to examine the natural trajectory of early sepsis more accurately, obviating being confounded by factors related to prolonged pre-ICU hospitalization. Approximately, 33.1% of the patient population was removed due to this criterion. Despite the exclusions, the distribution of major demographic and socioeconomic factors (age, gender, race, CCI, etc.) remained consistent with the full dataset, as shown in Supplemental Digital Content Table 1, http://links.lww.com/SHK/C380, confirming that these exclusions did not introduce significant bias.

Definition of clinical states and trajectory

To cluster patients, we first defined the following patient-centered states to capture a patient's clinical status during their ICU and hospital stay hourly. First, the clinical states were labeled based on admission and discharge status, with which we identified the following three states as: "discharge from hospital," denoting the patients have recovered and were discharged; "discharge from ICU," denoting the patients have improved and were transferred to a general hospital ward; and "death," denoting the patients who have succumbed to sepsis. Next, we further defined three states when patients were in the ICU: "In ICU," denoting the patients have been admitted into ICU but had no cardiovascular or respiratory dysfunction; "In ICU with single organ dysfunction," denoting the patients were in ICU with either cardiovascular or respiratory dysfunction; and "In ICU with both organ dysfunctions," denoting the patients were in ICU with both cardiovascular and respiratory dysfunctions.

Respiratory and cardiovascular dysfunctions were identified using real-time data on ventilation methods (high flow nasal cannula, noninvasive ventilation, invasive ventilation) and vasoactive medications (dobutamine, dopamine, epinephrine, norepinephrine, phenylephrine, and vasopressin), which were the most relevant variables available in our dataset. Other organ dysfunctions were excluded due to the low resolution of temporal data.

The first 2 weeks are crucial for monitoring sepsis progression and management, capturing peak organ dysfunction and intervention impacts. In our cohort, 81% of the patients were either discharged from hospital or deceased by 14 days. Therefore, we tagged data up to 14 days from the point of ICU admission. Because of data curation limitations, the EHR data for labeling organ dysfunctions in the ICU was limited to 7 days from admission. However, 89% of the patients were discharged from the ICU or deceased within 7 days and the ICU readmission rate is less than 10%, ensuring that no critical information regarding clinical states was missing for most patients. Overall, this high-resolution labeling provides a detailed view of the fluctuations in patient states and the interventions during this critical period.

Two-stage clustering method

In this study, clustering, an unsupervised machine learning approach, was applied to identify natural groupings within patient trajectories based on their clinical states over time. Clustering organizes data points into groups, or clusters, such that points within the same cluster are more similar to each other than to those in different clusters. This is particularly useful in clinical settings for discovering patterns without predefined labels, as it reveals inherent structures within the data.

Specifically, we employed a two-stage modeling approach. In the first stage, we used longitudinal K-means for initial broad clustering that captures overall temporal patterns due to its computational efficiency. In the second stage, DTW is used to refine these initial clusters by precisely aligning and identifying similar trends or shapes that may occur at different time points. This approach also allows us to account for both temporal (using K-means) and shape similarities (using DTW), thereby offering a comprehensive evaluation of similarity across individual time series. Additionally, validating clusters based on clinical expertise ensures that the identified patterns are meaningful in real-world healthcare settings. Moreover, our two-stage approach enhances the depth of analysis by allowing us to drill down into each of the four final clusters and examine finer subclusters within them. This detailed stratification reveals variations within the broader clusters, uncovering nuances that could be missed with a coarser analysis.

This approach leveraged the strengths of both methods, providing a comprehensive analysis of longitudinal data. The R packages kml and dtwclust were used. We also examined the association between identified clusters and baseline variables available at admission and during ICU stay, including demographics, SOFA score, APACHE III score, and Charlson Comorbidity Index (CCI). Data are presented as mean (standard deviation [SD]) or median (interquartile range [IQR]). Race is assessed using Fisher's exact test, gender using chi-square tests, and the remaining variables using ANOVA test. We initially tested across all identified clusters and then performed pairwise comparison with Bonferroni correction to control for false discovery. A significance level of 0.05 is used for assessment.

Segmented approach

As patients naturally have different lengths of stay, with 52% discharged from the hospital or deceased by day 3, their clinical trajectories exhibit unequal lengths. While padding is a common solution to address this issue in longitudinal K-means clustering, it may introduce bias, as clustering is based on distances measured across the entire 14-day trajectory. Therefore, we adopted a segmented approach in the longitudinal K-means clustering, segmenting patients based on their length

of hospital stay: 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, and 7 or more days, as clinical states became relatively stable from day 7 to day 14. This segmentation approach reduces the potential bias introduced by padding and allows for more accurate clustering based on the actual clinical trajectories.

Clustering metrics

For longitudinal K-means, an adaptation of the Calinski-Harabasz Index (22) proposed by Genolini *et al.* was used to select the best number of clusters, which is particularly applied in the context of longitudinal data clustering (23,24). For DTW, the silhouette index (25) was used to determine the optimal number of clusters to be selected. Clinical relevance was also considered when choosing the number of clusters for both longitudinal K-means and DTW. This ensures that the resulting clusters are statistically sound and meaningful in a clinical context.

RESULTS

Of 19,177 patients, 42% were female with a median age of 65 (IQR, 55–76) years, median APACHE III score of 70 (IQR, 56–87), hospital LOS of 7 (IQR, 4–12) days, and ICU LOS of 2 (IQR, 1–4) days. With the clinical states labeled, we examined the population flow from admission to day 14 by analyzing the state distribution at hour 0 (admission), hour 6, hour 12, hour 24, hour 72, day 7, day 10, and day 14, as shown in the Sankey plot in Figure 1. The first 6 h are the most critical for cardiorespiratory failure intervention, consistent with existing studies (26). By day 3, 62% of patients were discharged from the ICU. After 7 days, most patients were either discharged from the hospital or at least from the ICU, with much less frequent population change between states, indicating a steady state of patient flow.

Separate longitudinal K-means clustering models were applied to subpopulations segmented by their length of hospital stay, identifying 39 trajectories. These were further grouped using DTW based on similarities in local trends and patterns, identifying 7 clusters. Clinical expertise then consolidated these into four final clusters, requiring the combined clusters to share the same end point, that is, death or discharge alive, and allowing a range of days to be considered as delayed or slow progression, as shown in Figure 2. The study team then assigned the following clinically informed cluster names: fast recovery, slow recovery, fast decline, and delayed decline. Clinical states initially declined for all clusters in the first 6 h, indicated by a growing recipient of respiratory or cardiovascular interventions. Fast recovery, representing 27% of the population with a 3.5% death rate and a median hospital LOS of 3 (IQR, 2–15) days, typically showed immediate improvement after 6 h and a quick discharge within 3 days. In contrast, slow recovery, covering 62% of the population with a 3.6% death rate and a median hospital LOS of 8 (IQR, 6–13) days, exhibited an initial decline often followed by gradual improvement. Unlike fast recovery, the recovery trajectory for the slow recovery population is flatter and involves prolonged ICU stay. Fast decline, comprising 3% of the population with a 99.7% death rate and a median hospital LOS of 1 (IQR, 0–1) day, showed immediate deterioration shortly after admission. In contrast, delayed decline, representing 7% of the population with a 97.9% death rate and a median hospital LOS of 5 (IQR, 3–8) days, initially stabilized for a period of time before quickly deteriorating. Unlike the fast decline population, which showed no improvement after intervention and progressed directly to death, the delayed decline group had a brief period of stabilization or even some improvements before their condition worsened. The period before deterioration varies depending on the ICU LOS, but shows similar shape in clinical trajectory.

Table 1 summarizes the distributions for demographics, CCI, SOFA scores, and APACHE III score across the identified clusters. Major determinants for different clinical trajectories included CCI, 24-h SOFA scores, day 3 SOFA scores, and APACHE III scores, which aligns with the observed differences in mortality and hospital and ICU LOS across clusters. Median 24-h SOFA score was highest in the "fast decline" cluster (11) and lowest in the "fast recovery" cluster (5). Similar patterns were observed for day 2 and day 3 SOFA scores. Median 24-h APACHE III score was highest in the "fast decline" cluster (109) and lowest in the "fast recovery" cluster (63). Age and BMI were significant demographic factors. Median age was highest in the "delayed decline" cluster (70 years) and lowest in the "fast recovery" cluster (64 years). While gender and race were initially significant at the 0.05 level when tested across the four clusters, further pairwise tests (i.e., comparing each pair of categories of the cluster, for instance, slow recovery *versus* fast recovery, as one pair) indicated no significant differences in race or gender between clusters, suggesting that the initial significance may be due to the large sample size rather than actual differences.

The additional comparison results related to comorbidities are summarized in Table 2, http://links.lww.com/SHK/C380 in the supplemental material.

DISCUSSION

Clinically meaningful cluster discovery

Using an unsupervised machine learning two-stage clustering approach, we successfully identified four distinct clinical trajectories among patients with sepsis during their hospital stay, with a focus on cardiorespiratory failure and a high time resolution of

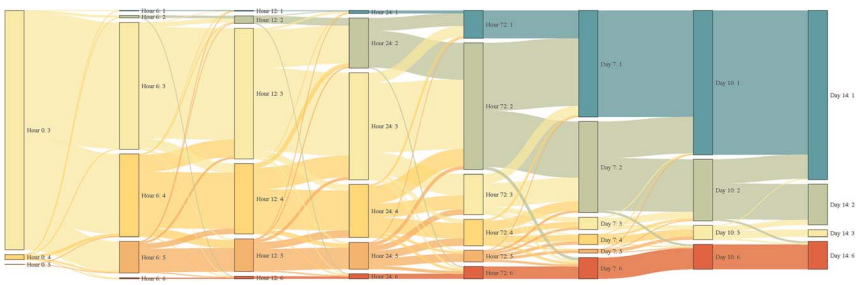

Fig. 1. **Sankey plot illustrating the population flow between clinical states from admission (hour 0) to 14 days.** The states are defined as follows: 1 – Discharge from hospital, 2 – Discharge from ICU, 3 – In ICU without cardiovascular or respiratory dysfunction, 4 – In ICU with either cardiovascular or respiratory dysfunction, 5 – In ICU with both cardiovascular and respiratory dysfunction, 6 – Death. The link represents the number of changes between adjacent time points. The width of the flow represents the population size leaving that state.

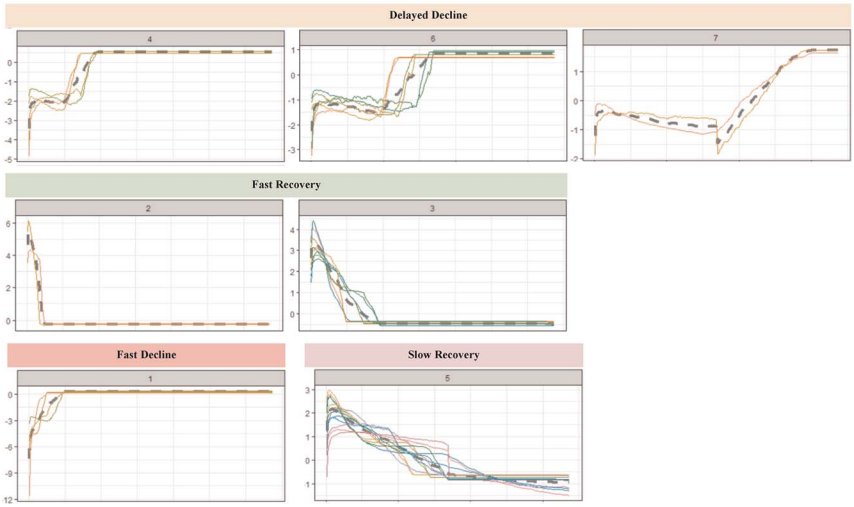

Fig. 2. **Four clusters including delayed decline (upper panel), fast recovery (middle panel), fast decline and slow recovery (lower panel) are identified.** The Y-axis represents the clinical states that are z-normalized by trajectory to highlight the direction and magnitude of the change instead of the specific shifts between different states. A higher value indicates a worse clinical state, while a flat line indicates the absorbative states like decease or discharge alive. The X-axis represents the hours since ICU admission (hour 0) up to 14 days. Each subplot corresponds to a cluster identified by DTW. Colored lines within each subplot represent trajectories identified by longitudinal K-means clustering. Dashed line indicates the overall trajectory identified by DTW within each cluster based on the centroids calculated. DTW, dynamic time warping; ICU, intensive care unit.

1 h. This work is the first that looks into extended hospital stay for up to 14 days. While this work focused on the critically ill patients with sepsis, once validated, this can be utilized for other disease states commonly seen in the intensive care units such as gastrointestinal bleeding and acute respiratory distress syndrome among others.

Existing research on trajectory clustering has typically focused on short-duration longitudinal data, ranging from 8 to 72 h (19). However, for sepsis, the first 6 h often involve intensive interventions to stabilize the patient's vitals and overall condition. Consequently, the trajectories based on vitals or organ dysfunctions are heavily influenced by these interventions, masking the underlying differences crucial for accurate phenotyping. Even extending the observation window to 72 h may not be sufficient, as a significant proportion of patients (32% in our cohort) remain in the ICU beyond this period. These patients exhibit vastly different clinical trajectories after 72 h, as shown in our results. For instance, both the slow recovery and delayed decline groups have patients staying in the ICU for more than 72 h. While the slow recovery group gradually improves, the delayed decline group could rapidly deteriorate afterwards. Without extending the observation window to a longer time frame, for example, 14 days, distinguishing between these two groups would be challenging. Nonetheless, identifying the delayed decline group is crucial for clinical practice to find potential solutions to prevent deterioration after stabilization. Our method, with an extended observation period

Table 1. **Summary statistics by clusters, with median (IQR) for continuous variables and distribution for categorical variables**

	Fast decline	Fast recovery	Slow recovery	Delayed decline	*P*
n (%)	665 (3.5%)	5,156 (26.9%)	11,973 (62.4%)	1,383 (7.2%)	
Mortality (%)	663 (99.7%)	181 (3.5%)	431 (3.6%)	1,354 (97.9%)	
Hospital LOS, median (IQR)	1 (0, 1)	3 (2, 15)	8 (6, 13)	5 (3, 8)	
ICU LOS, median (IQR)	1 (0, 1)	1 (1. 2)	3 (1, 5)	3 (2, 6)	
CCI, median (IQR)	7 (4, 9)	6 (4, 9)	6 (4, 9)	7 (4, 9)	<0.001
24-h SOFA score, median (IQR)	11 (8. 14)	5 (3, 7)	6 (4, 9)	8 (6, 11)	<0.001
Day 2 SOFA score, median (IQR)	8 (4, 11)	2 (1, 4)	4 (2, 7)	7 (4, 11)	<0.001
Day 3 SOFA score, median (IQR)	8 (4. 11)	2 (1, 4)	3 (1, 6)	7 (4, 10)	<0.001
24-Hour APACHE III, median (IQR)	109 (87, 135)	63 (51, 77)	70 (57, 86)	88 (72, 108)	<0.001
Age, median (IQR)	69 (58, 81)	64 (53, 74)	66 (55, 76)	70 (58, 80)	<0.001
BMI, median (IQR)	27 (23, 33)	28 (24, 33)	29 (24, 34)	27 (23,33)	<0.001
Gender					0.04
Female (%)	298 (45%)	2083 (40%)	5,052 (42%)	561 (41%)	
Male (%)	367 (55%)	3,073 (60%)	6,921 (58%)	822 (59%)	
Race					0.02
American Indian/Hawaiian Native/Pacific Islander (%)	2 (0.3%)	45 (0.9%)	81 (0.7%)	9 (0.7%)	
Asian (%)	9 (1.4%)	77 (1.5%)	187 (1.6%)	15 (1.1%)	
Black/African (%)	8 (1.2%)	123 (2.4%)	236 (2.0%)	20 (1.5%)	
Other (%)	39 (5.9%)	261 (5.1%)	508 (4.2%)	74 (5.4%)	
White (%)	607 (91.3%)	4,650 (90.2%)	10,961 (91.5%)	1,265 (91.5%)	

For *P* values, race is assessed using Fisher's exact test, gender using chi-square tests, and the remaining variables using ANOVA test.

of up to 14 days, provides a comprehensive and dynamic view of sepsis progression, capturing the intricacies of cardiorespiratory failure and hospital discharge status.

While mortality and LOS are critical outcomes, they provide a limited view of the complex and dynamic progression of sepsis, and focusing solely on mortality or LOS ignores the intermediate states that patients go through, which are critical for understanding the progression and management of sepsis. Some studies clustered patients with sepsis using SOFA score (19) which derived similar concepts; however, the quantitative statistics of these four clusters are different from ours. In addition, the time resolution is crude and the SOFA score alone cannot fully reflect the intermediate states, including periods of stabilization, temporary improvements, and setbacks that can provide valuable insights into the effectiveness of interventions and the patient's response to treatment. Understanding these stages is essential for developing targeted therapies and identifying actionable time points in the clinical trajectories that could alter and hopefully improve patient outcomes. This study's approach offers a holistic and dynamic view of patient progression and management, capturing the full spectrum of clinical changes from ICU admission to discharge or death.

Methodological strengths (innovative advanced clustering approach)

Comparing to existing studies, our study is featured by the use of high-resolution EHR data to create detailed clinical trajectories, providing a granular view of patient progression. Existing studies have explored K-Means clustering, latent class analysis, latent profile analysis. DTW, and group-based trajectory models (5–7,12,16,17,19,27,28). The methodological strength of this study lies in its two-stage clustering approach. This method is more scalable than DTW alone and captures local trends or shapes that longitudinal K-Means alone cannot. By combining these methods, we achieved a balance of computational efficiency and detailed pattern recognition, providing a robust framework for analyzing complex longitudinal data. High-resolution, multistage clustering analysis performed in our study offers a more comprehensive and actionable understanding of sepsis progression.

Foundation for digital twin development

The clinical trajectories derived can be used to enhance the fidelity of patient modeling and further inform digital twin development of the ICU system for operational management. Virtual patients can be created to mirror the recovery patterns using the clinical trajectories identified in this study, allowing for the conceptualization of complex patient conditions into simplified, easy-to-interpret models. As a first-step analysis, we explored what phenotypes exist in the patient population. Then, different models, based on patients' specific phenotype, can be derived. As an example, multistate models based on the clinical states (i.e., death, discharge, ICU no organ dysfunction, single organ dysfunction, multiple organ dysfunction) that define the four trajectories can be developed (29). These models can predict the clinical state of a patient at a moment during their ICU stay. By incorporating these patient models in the ICU digital twin, we can enhance our ability to predict ICU census, and further investigate the ICU clinician workload based on patients' organ system dysfunction states. Current ICU simulation typically ignores the patient-specific characteristics and limits its capability to explore specific potential scenarios. By mapping each trajectory to different levels of care needs, we will be able to evaluate various scenarios, such as a shift in the volume and composition of patients following different trajectories. For instance, during the COVID pandemic, the percentage of patients following the rapid decline trajectory will increase, whereas the population of the low-risk monitoring patients might decrease. The impact of this change on ICU census and clinician utilization can be evaluated accordingly. This method is not only valuable for sepsis management but is also transferable to other domains, enabling personalized care and more efficient hospital operations across a wide range of medical conditions.

Limitations

This is a single center study using data exclusively from Mayo Clinic collected during 2011–2018. However, multiple ICUs, including medical, surgical, multispecialty and transplant, cardiovascular medicine and surgery ICU data were included to enhance the generalizability of the results. To suggest or prove any external generalizability the model will have to be prospectively and externally validated. In addition, this work serves as a model development and derivation. This methodology can be generalized to other institutions and other patient populations. The study is limited by the availability of detailed organ dysfunction data only for the first 7 days of ICU stay. In addition, this study only included cardiorespiratory dysfunction for the trajectory tagging. However, the clinical management for patents with sepsis has remained fairly unchanged during this period. Additionally, clustering approaches, including ours, are inherently heuristic and can be influenced by factors such as initialization methods, randomness, normalization methods, distance measures, or data splits. Despite these variations, our analysis showed that the core qualitative structure of the four identified clusters—fast recovery, slow recovery, fast decline, and delayed decline—remained stable, confirming the reliability of our clustering approach. This consistency suggests that the overarching patterns are robust, even though individual cluster assignments may shift due to methodological choices.

Future work

Future studies may include other organ dysfunctions to further refine clinical trajectory. Further validation on larger and more diverse cohorts will ensure the external validity and generalizability of the findings. Moreover, the clinical trajectories were derived based on the aggregated organ dysfunction states. This level of abstraction is sufficient for supporting strategic or tactic level ICU management decisions. The model's clinical utility could be enhanced by incorporating a broader range of real-time physiological parameters. This would be ideal for the patient-level modeling for clinical decisions. For instance, the digital twin built upon this framework can simulate different treatments that possibly alter patient clinical trajectories to support informed decision making. Lastly, the likelihood that the patient belongs to a specific trajectory is in part determined by the baseline chronic health

status and specific comorbidities. Further study is required to better understand the mechanisms of these associations.

CONCLUSIONS

This study comprehensively depicts sepsis progression through detailed clinical trajectories, offering valuable insights to facilitate more robust prediction models and digital twin decision support tools for both shared decision making and resource planning. By identifying distinct clinical trajectories through clustering, these models provide a nuanced understanding of patient progression, moving beyond static assessments toward dynamic, real-time analyses. Specifically, the clustering results serve as the foundation for creating digital twin models that mirror patient-specific disease evolution, allowing clinicians to better anticipate patient needs, manage patient communications, allocate resources effectively, and improve overall care quality for patients with sepsis.

ACKNOWLEDGMENTS

Authors' contributions: XZ is the guarantor of the content of the manuscript, including the data and analysis. GYH, XZ, AL, PJS, and OG contributed to the conceptualization of the study. GYH and XZ contributed to the data analysis and methodology design. GYH and XZ contributed to the drafting of the first version of the manuscript. All authors participated in the interpretation of the data, critically reviewed the manuscript, provided final approval for submission, and take responsibility for the accuracy and integrity of the work

REFERENCES

1. Rhee C, Dantes R, Epstein L, et al. Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009-2014. *JAMA*. 2017;318(13):1241–1249.
2. Bauer M, Gerlach H, Vogelmann T, et al. Mortality in sepsis and septic shock in Europe, North America and Australia between 2009 and 2019—results from a systematic review and meta-analysis. *Crit Care*. 2020;24(1):239.
3. Lal A, Li G, Cubro E, et al. Development and verification of a digital twin patient model to predict specific treatment response during the first 24 hours of sepsis. *Crit Care Explor*. 2020;2(11):E0249.
4. Lone NI, Walsh TS. Impact of intensive care unit organ failures on mortality during the five years after a critical illness. *Am J Respir Crit Care Med*. 2012;186(7):640–647.
5. Seymour CW, Kennedy JN, Wang S, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. *JAMA*. 2019;321(20):2003–2017.
6. Bhavani SV, Semler M, Qian ET, et al. Development and validation of novel sepsis subphenotypes using trajectories of vital signs. *Intensive Care Med*. 2022;48(11):1582–1592.
7. Yehya N, Fitzgerald JC, Hayes K, et al. Temperature trajectory sub-phenotypes and the Immuno-inflammatory response in pediatric sepsis. *Shock*. 2022;57(5):645–651.
8. Trevena W, Zhong X, Lal A, et al. Model-driven engineering for digital twins: a graph model-based patient simulation application. *Front Physiol*. 2024;15:1424931.
9. Rovati L, Gary PJ, Cubro E, et al. Development and usability testing of a patient digital twin for critical care education: a mixed methods study. *Front Med (Lausanne)*. 2023;10:1336897.
10. Zhong X, Babaie Sarijaloo F, Prakash A, et al. A multidisciplinary approach to the development of digital twin models of critical care delivery in intensive care units. *Int J Prod Res*. 2022;60(13):4197–4213.
11. Zhong X, Abrol Neena S, Hou GY, et al. A hybrid simulation approach for modeling critical care delivery in ICU. In *Proceedings of the 2024 Winter Simulation Conference*. Orlando, FL; 2024.
12. Demerle KM, Angus DC, Baillie JK, et al. Sepsis subclasses: a framework for development and interpretation*. *Crit Care Med*. 2021;49(5):748–759.
13. Davenport EE, Burnham KL, Radhakrishnan J, et al. Genomic landscape of the individual host response and outcomes in sepsis: a prospective cohort study. *Lancet Respir Med*. 2016;4(4):259–271.
14. Scicluna BP, van Vught LA, Zwinderman AH, et al, MARS consortium. Classification of patients with sepsis according to blood genomic endotype: A Prospective Cohort Study. *Lancet Respir Med*. 2017;5(10):816–826.
15. Sweeney TE, Azad TD, Donato M, et al. Unsupervised analysis of transcriptomics in bacterial sepsis across multiple datasets reveals three robust clusters. *Crit Care Med*. 2018;46(6):915–925.
16. Ding YJ, Luo Z, Szeto M, Luo Y, Sanchez-Pinto LN. Pediatric sepsis phenotyping using vital sign trajectories. In: *Proceedings* - 2023 *2023 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2023*. Institute of Electrical and Electronics Engineers Inc; 2023:4297–4303. doi:10.1109/BIBM58861.2023.10386019
17. Bhavani SV, Carey KA, Gilbert ER, et al. Identifying novel sepsis subphenotypes using temperature trajectories. *Am J Respir Crit Care Med*. 2019;200(3):327–335.
18. Bhavani SV, Xiong L, Pius A, et al. Comparison of time series clustering methods for identifying novel subphenotypes of patients with infection. *J Am Med Inform Assoc*. 2023;30(6):1158–1166.
19. Xu Z, Mao C, Su C, et al. Sepsis subphenotyping based on organ dysfunction trajectory. *Crit Care*. 2022;26(1):197.
20. Singer M, Deutschman CS, Seymour C, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA*. 2016;315(8):801–810.
21. Lal A, Rayes H, O'Horo JC, et al. Septic shock definitions and associated outcomes in blood culture positive critically ill patients. *Ann Transl Med*. 2023;11(5):192–192.
22. Caliński T, Harabasz J. A dendrite method foe cluster analysis. *Communications in Statistics*. 1974;3(1):1–27.
23. Genolini C, Alacoque X, 1027 IU, Toulouse C, Sentenac M, Arnaud C. *Kml and Kml3d: R Packages to Cluster Longitudinal Data*. Vol 65.; 2015. Available at: http://www.jstatsoft.org/. Accessed January 17, 2025.
24. Genolini C, Falissard B. *KmL: A Package To Cluster Longitudinal Data*.
25. Rousseeuw PJ. *Silhouettes: a graphical aid to the interpretation and validation of cluster analysis*. *J Comput Appl Math*. 1987;20:53–65
26. Gauer RL. *Early recognition and management of sepsis in adults: the first six hours* Am Fam PhysicianVol 88.; 2013: 44-53. www.aafp.org/afp.
27. Hao C, Hao R, Zhao H, et al. Identification and validation of sepsis subphenotypes using time-series data. *Heliyon*. 2024;10(7):e28520.
28. Zhang Z, Zhang G, Goyal H, et al. Identification of subclasses of sepsis that showed different clinical outcomes and responses to amount of fluid resuscitation: a latent profile analysis. *Crit Care*. 2018;22(1):347.
29. Klein Klouwenberg PMC, Spitoni C, Van Der Poll T, et al. Predicting the clinical trajectory in critically ill patients with sepsis: a cohort study. *Crit Care*. 2019;23(1):408.