Sepsis Subphenotyping Based on Organ Dysfunction Trajectory

Zhenxing Xu, PhD[1], Chengsheng Mao, PhD[2], Chang Su, PhD[1], Ilias Siempos, MD[3,4], Lisa K Torres, MD[3,4], Di Pan, DO[4], Yuan Luo[*], PhD[2], Edward J Schenck[*], MD[3], Fei Wang[*], PhD[1]

[1]Division of Health Informatics, Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, USA; [2]Division of Health and Biomedical Informatics, Department of Preventive Medicine Center for Health Information Partnerships, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA; [3]Division of Pulmonary and Critical Care Medicine, NewYork-Presbyterian Hospital/Weill Cornell Medical Center, New York, NY, USA; [4]Weill Cornell Medical College, Weill Cornell Medicine, New York, NY, USA; *Corresponding authors.

Corresponding author
Yuan Luo: yuan.luo@northwestern.edu
Edward J Schenck: ejs9005@med.cornell.edu
Fei Wang: few2001@med.cornell.edu

Abstract (250 words)

Purpose: Although organ dysfunction is a defining element of sepsis, its trajectory is not well studied. We sought to identify whether there are distinct Sequential Organ Failure Assessment (SOFA) score trajectory-based subphenotypes in sepsis.

Methods: We created 72-hour SOFA score trajectories in patients with sepsis from two diverse intensive care unit (ICU) cohorts. We then used Dynamic Time Warping (DTW) to compute patient similarities to capture evolving heterogeneous sequences and establish similarities between groups with distinct trajectories. Hierarchical agglomerative clustering (HAC) was utilized to identify subphenotypes based on SOFA trajectory similarities. Patient characteristics were compared between subphenotypes and a random forest model was developed to predict subphenotype membership, within 6 hours of ICU arrival. The model was then tested on the validation cohort.

Results: A total of 4,678 and 3,665 unique sepsis patients were included in development and validation cohorts. In the development cohort, four subphenotypes of organ dysfunction were identified: Rapidly Worsening (n=612, 13.08%), Delayed Worsening (n=960, 20.52%), Rapidly Improving (n=1,932, 41.3%) and Delayed Improving (n=1174, 25.1%). In-hospital mortality for patients within different subphenotypes demonstrated distinct patterns over time. Similar subphenotypes and their associated outcome patterns were replicated in the multicenter validation cohort.

Conclusion: Four novel, clinically-defined, trajectory-based sepsis subphenotypes were identified and validated. Trajectory based subphenotyping is useful for describing the natural history of sepsis in the ICU. Understanding the pathophysiology of these differential trajectories may reveal unanticipated therapeutic targets for patients with sepsis and identify more precise populations and endpoints for the predictive enrichment of clinical trials.

Keywords: Sepsis, Subphenotype, SOFA, Precision medicine, Dynamic time warping

Main text: (2,993 words)

Introduction

Sepsis is defined as a dysregulated immunological response to infection that results in acute organ dysfunction [1]. The morbidity and mortality of sepsis remain high despite decades of research [2] and numerous failed clinical trials [3]. Recent research has highlighted that sepsis is a complex and heterogeneous syndrome, which includes a multidimensional array of clinical and biological features [4]. Identifying rigorous sepsis subphenotypes that present with similar prognostic markers and pathophysiologic features has the potential to improve therapy [5-8].

Most prior sepsis subphenotyping studies use static measurements available soon after admission to the emergency department or intensive care unit (ICU) to characterize patients [4, 9-11]. However, due to the stochastic nature of the initiation of infection and variable presentation to health care after developing symptoms, static assessments of sepsis subphenotypes may be incomplete, given the dynamic nature of the immune response and evolution of organ failure in sepsis [12].

More recently, dynamic trajectories of patient temperature have been identified in sepsis. The differential pattern of temperature change may represent a varied underlying inflammatory response to infection [1]. Whether multisystem organ failure develops distinct phenotypic patterns over time in sepsis remains largely unexplored. Identifying distinct organ dysfunction trajectories in sepsis can refine our understanding of the natural history of sepsis in the ICU in response to standard of care treatment and define patterns of disease that may benefit from novel therapeutic strategies [13].

In this study, we examined whether there are distinct subphenotypes of organ dysfunction trajectory over the first three days of an ICU admission with sepsis. We used longitudinal SOFA scores and Dynamic Time Warping (DTW) [14] to evaluate patients' organ dysfunction trajectory similarities, as a surrogate for disease evolution. Clinical and biologic features including demographics, comorbidities, and inflammatory variables were used to characterize these subphenotypes at baseline. We then explored whether baseline features could predict these subphenotypes soon after ICU admission.

Methods

Overview

We did a cohort study on datasets from two ICU cohorts that contained granular patient level data from a total of twelve hospitals in the United States, whose overall workflow is illustrated in Fig. 1. Our goal was to derive sepsis subphenotypes of patients in ICU according to their SOFA organ dysfunction trajectories using Hierarchical Agglomerative Clustering (HAC). We then characterized these subphenotypes using comprehensive patient information including demographics, comorbidities, use of mechanical ventilation, type of ICU unit, admission source, organ source of sepsis, and examined their associated clinical outcomes as well as clinical biomarkers. We further built a random forest model to predict the derived subphenotypes from baseline patient clinical characteristics. To ensure reproducibility, the same analysis pipeline was conducted on the validation cohort.

Definition of sepsis and study population

The development cohort was from the Medical Information Mart for Intensive Care III (MIMIC-III) database [15], derived from Beth Israel Deaconess Medical Center (BIDMC), from admissions dating from 2001-2012. BIDMC is a teaching hospital of Harvard Medical School in Boston, Massachusetts with 673 licensed beds, including 493 medical/surgical beds, 77 critical care beds, and 62 OB/GYN beds. The validation cohort was from the health system of Northwestern Medicine Enterprise Data Warehouse (NMEDW) [16], a network of eleven hospitals located in northern Illinois with 2,554 beds in total, with ICU admissions dating from 2012-2019. More details about both cohorts are shown in Supplemental Appendix 1. The inclusion-exclusion cascade for the patients in our study are shown in Supplemental Fig. S1, where Sepsis-3 criteria are defined as in Singer et al. [17].

SOFA score computation, model descriptions, and statistics analysis

The SOFA score was derived from six organ-specific subscores including respiration, coagulation, liver, cardiovascular, central nervous system (CNS), renal [13]. In our study, SOFA scores were derived every 6 hours within the first 72 hours of ICU admission. For each 6-hour period, the worst variable value was used to compute the SOFA subscores. Missing values were imputed using last observation carried forward (LOCF) or first observation carried backward (FOCB) [13].

After obtaining the SOFA trajectories of patients, we used DTW [14] to evaluate pairwise trajectory similarities, based on which the subphenotypes were identified through HAC [18]. McClain index [19] was used to determine the optimal numbers of subphenotypes.

After the subphenotypes were obtained, we used statistical testing to examine the associations between different clinical variables and biomarkers with the subphenotypes. Detailed list of clinical variables and biomarkers, as well as the concrete statistical testing techniques adopted, are shown in Supplemental Appendix 2.

Subphenotype prediction

We trained a random forest model to predict the derived subphenotypes from the baseline patient clinical characteristics within the first 6 hours after ICU admission, with the goal of examine whether the trajectory subphenotypes could be predicted early. Candidate predictors included demographics, comorbidities, SOFA subscores, lab tests, and vital signs. Overall prediction accuracy, along with the precision and recall each class are reported with the one-vs-rest scheme. Predictor contributions were evaluated with the Shapley additive explanations (SHAP) strategy [20].

Results

Basic cohort characteristics

For the development cohort (n=4,678), the median age was 65.93 years (IQR [53.71-77.88]). It included 2,625 male and 3367 white (71.98%). The overall in-hospital mortality rate was 10.86%, and the median ICU length-of-stay was 2.76 days (Interquartile Range (IQR) [1.59-5.57]). There were 1,893 patients (40.47%) treated with mechanical ventilation during the first three days. The mean baseline SOFA score obtained from the first 6 hours was 4.96 (Standard Deviation (SD): 2.80). Most of the patients (2,611, 55.81%) were in the medical intensive care unit (MICU). The demographic distribution of the validation cohort (n=3,665), is similar to the development cohort. The overall in-hospital mortality rate of 14.02% and the median length-of-stay was 3.84 days (IQR [1.98-7.94]). There were 1,524 patients (41.58%) that needed mechanical ventilation in the first three days. The mean baseline SOFA score was 5.68 (SD:2.81).

Comparisons between Survivors and Nonsurvivors

Nonsurvivors in development cohort were older than survivors, with a median age of 71.46 years (IQR, [59.87-80.93]) compared with 65.22 years for survivors (IQR, [53.18-77.35], p-value < 0.001). Nonsurvivors had higher comorbidity burden with a median Elixhauser index score [21] 7.0 (IQR [2.0-12.0]). Median ICU length-of-stay for nonsurvivors was 3.95 days (IQR [1.88-7.7]), and the rate of mechanical ventilation during the first three days was 59.84%. Nonsurvivors had higher baseline SOFA scores, with a mean value 7.10 (SD: 3.69). More nonsurvivors were admitted in MICU. More details about differences between survivors and nonsurvivors were shown in Supplemental Table S1, with similar statistics in validation cohort were shown in Supplemental Table S2.

SOFA trajectory and the derived subphenotypes

Based on the pairwise patients' SOFA trajectory similarity matrix obtained from DTW, we generated clustermaps as in Supplemental Fig. S2, where four distinct clusters were identified through optimizing the McClain index [19].

Specifically, in the development cohort (Fig. 2(a)), Rapidly Worsening (n=612, 13.08%) was characterized by continuously increased SOFA scores from a mean (SD) of 4.52 (2.80) at admission to more than 7 at 72 hours. This subphenotype had the fewest patients. Delayed Worsening (n=960, 20.52%) was characterized by decreased SOFA scores within the first 48

hours from a mean (SD) of 5.23 (2.73) at baseline to 3.65 (2.83), followed by an increase over the last 24 hours. Rapidly Improving (n=1,932, 41.3%) was characterized by a consistent continuous improvement in SOFA scores from a mean (SD) of 5.54 (2.92) in baseline to less than 3. This was the most common and had the highest SOFA score at baseline. Delayed Improving (n=1174, 25.1%) was characterized by an increase and then a gradual decrease in SOFA score over 72 hours. It had the lowest SOFA score at baseline with mean (SD) 4.02 (2.36). Similar trajectory trends were obtained in the validation cohort (Fig. 2(b)) and the detailed analysis was provided in Supplemental Appendix 3. Individual SOFA subscore trajectories for each subphenotype were shown in Supplemental Fig. S3 and Supplemental Fig. S4.

Patient characteristics comparisons across subphenotypes

Patient characteristics differed across the derived subphenotypes (see Table 1, Fig. 2(c) and (d), and Fig. 3). Specifically, Rapidly Worsening patients had the highest rates of mechanical ventilation (46.41%), the highest median Elixhauser comorbidity burden value of 5 (IQR [0-10]) but the lowest baseline SOFA score compared to the other subphenotypes. They had the worst mortality (Fig. 2(c) 28.27%, p-value<0.001) and longer length-of-stay (Table 1, 2.88 days, p-value<0.001). Rapidly Improving patients had the lowest rate of mortality (Fig. 2(c) 5.54%) and mechanical ventilation (37.94%), and the shortest length-of-stay (2.42 days). It had the highest proportion of patients meeting criteria for septic shock (15.48%, p-value=0.002). Delayed Improving and Delayed Worsening patients had lower rates of mortality (10.73%, 10.63%) and mechanical ventilation (42.50%, 39.27%) than the Rapidly Worsening subphenotype. The median age of the four subphenotypes were similar in the development cohort. Male patients were more common in all subphenotypes. More information was shown in Table 1. In addition, clinical biomarkers were different among subphenotypes. Chord diagrams (Fig. 3) showed these differences of subphenotypes in terms of abnormal clinical variables. The Rapidly Worsening group was more likely to have patients with abnormal cardiovascular biomarkers (bicarbonate, troponin T or I, lactate) and hematologic (such as hemoglobin, INR, platelet, glucose, RDW). Patients in this subphenotype had a higher chronic comorbidity burden and had abnormal SOFA subscores including respiration, coagulation and liver. The Rapidly Improving patients were more likely to have abnormal inflammatory lab values (temperature, WBC, bands, CRP, albumin, lymphocyte percent) and abnormal cardiovascular, CNS and renal SOFA subscores. There was a lower chronic comorbidity burden in this subphenotype. Delayed Worsening group had more abnormal hematologic and respiration, coagulation, CNS, and SOFA renal subscores. Abnormal respiration, coagulation, cardiovascular SOFA subscores were strongly associated with Delayed Improving. More information was shown in Supplemental Appendix 4 and Appendix 5.

Subphenotype prediction

We trained random forest models for predicting the four subphenotypes according to early stage characteristics. Overall, for four subphenotypes, the prediction models obtained the accuracy of 0.75 (95% Confidence Interval [CI], [0.73, 0.76]) on development cohort and

0.79 (95% CI [0.77, 0.80]) on validation cohort. The precision and recall for each subphenotype on both cohorts are shown in Supplemental Table S6. Predictor contributions on both cohorts were shown in Fig. 4 and Fig. S5, which demonstrated different patterns when predicting different subphenotypes. For example, RDW, creatinine, bicarbonate and BUN contributes more for predicting the Rapidly Improving group, while platelet, INR, AST and lactate contributed more to the prediction of the Rapidly Worsening group.

Discussion

We reported four sepsis subphenotypes based on dynamic organ dysfunction trajectories using a novel data-driven methodology. DTW was used to calculate patients' SOFA trajectory similarities because of its capability of capturing heterogeneous evolution among the temporal sequences more robustly, based on which HAC was leveraged to identify patient groups with similar trajectories. Four progression subphenotypes were identified, Rapidly Worsening, Delayed Worsening, Rapidly Improving, and Delayed Improving. Patients in the Rapidly Worsening subphenotype had progressive organ dysfunction over the time period. The two Delayed groups had unstable organ dysfunction over the study period and the Rapidly Improving group had the highest admission organ dysfunction but quickly improved. Outcomes followed SOFA trajectory across each subphenotype irrespective of traditional septic shock categories, Rapidly Improving had the best outcomes and Rapidly Worsening the highest mortality. The Delayed Improving and Delayed Worsening had intermediate outcomes, including mortality, and length of stay.

The potential for distinct pathophysiologic etiologies for the differential trajectories is supported by the differential patterns of organ dysfunction, vital signs, inflammatory, hematologic, and cardiovascular variables at admission to the ICU. As shown in Fig. 3, and Supplemental Fig. S6 and S7, there were different variables that were associated with the groups over the course of the study. For example, those patients of Rapidly Improving were more likely to have more abnormal inflammatory markers (such as WBC, bands, CRP, albumin, temperature, lymphocyte) and more abnormal values on cardiovascular, and CNS subscores. There was a lower comorbidity score in patients with this subphenotype, which suggests that sepsis outcomes may be more dependent on underlying illness. The Rapidly Worsening patients had more comorbidities and distinct derangements in clinical variables associated with metabolic acidosis and hypoperfusion, e.g. a low bicarbonate and higher lactate, and disseminated intravascular coagulation, e.g. low platelets and a higher INR and respiratory failure. Both of the Delayed subphenotypes had a less specific group of variables associated with group membership, including inflammatory, hepatic, hematologic and pulmonary associated with Delayed Improvement and hematologic, cardiovascular and renal variables associated with Delayed Worsening.

Importantly, we built a multivariable prediction model for the identified trajectory subphenotypes from patient baseline characteristics and early-stage clinical features. Models were built on the first 6 hours after being admitted to ICU. Several interesting findings were obtained. For example, (1) A high comorbidity score tended to predict the subphenotypes of Rapidly Worsening because patients with high comorbidity score had multiple diseases and

were more likely to present worse organ dysfunction in ICU. (2) The roles of lab tests and vital signs were different on prediction. For example, low Platelets had a positive impact on the Rapidly Worsening prediction and high Platelets had a positive impact on the Rapidly Improving prediction. The high bands and bilirubin tended to predict the subphenotype of Delayed Improving. The high values of lactate and INR tended to predict the subphenotype of Rapidly Worsening. (3) Some features such as BMI and lymphocyte count did not contribute to discriminating the subphenotypes significantly.

Our manuscript complements and adds to other recent study of sepsis subphenotypes. For example, Seymour et al. [4] and Knox et al. [9] each identified four subphenotypes that were associated with organ dysfunction patterns and clinical outcomes in patients with sepsis using a panel of baseline clinical variables. There is some overlap in our high risk groups, notably both include liver injury and shock. However, our work demonstrates that the difference in outcome in this group is due to progressive non-resolving organ dysfunction that calls for novel treatments. Bhavani et al. [1] identified novel longitudinal temperature trajectories to identify four sepsis subphenotypes, with significant variability in inflammatory markers and outcomes, highlighting the potential for novel immune signatures to be uncovered through trajectory analysis.

A major strength of this analysis is that we have identified time-dependent progression patterns that may be related to the differential response of specific organ dysfunction to standard of care interventions. For example, the Rapidly Improving group had cardiovascular and respiratory failure at admission that resolved over 72 hours. The Rapidly Worsening groups developed multisystem organ failure including visceral organ dysfunction, specifically liver failure in addition to cardiovascular and respiratory failure. These differential patterns suggest varying time-dependent, treatment responsive organ dysfunction pathophysiology in sepsis. The cardiovascular and respiratory subscores are driven by the vasopressor dose and PaO2/FiO2 respectively, which may respond to therapeutic interventions such as corticosteroids, volume resuscitation, and the application of PEEP or therapeutic suctioning. [22]. However, as demonstrated by our analysis sepsis-related renal and liver failure may be less modifiable with our current therapeutic strategies over the past twenty years [23, 24]. Our study highlights that patterns of organ dysfunction in patients with sepsis are Rapidly Improving, Rapidly Worsening and Delayed. Each of these patterns may be due to a different pathophysiology and benefit from different treatments in the future.

It deserves noting that our Rapidly Improving patients had better outcomes across all patients studied, but still represented 21% and 36% of all deaths in our derivation and validation cohorts respectively, despite an overall 5% and 10% in-hospital mortality. This low mortality rate has implications for powering clinical trials [25] and further research is needed to understand the cause of death in patients with rapidly improving organ dysfunction in sepsis. The Rapidly Worsening subphenotype was the least common and may represent patients with our classical understanding of hyperinflammatory septic shock [26]. More recent evidence suggests that the pathophysiology of early, progressive organ dysfunction in our Rapidly

Worsening patients may be due to over exuberant activation of necroinflammatory cell death pathways in multiple organs, highlighting the need for novel treatment strategies [27-29]. The Delayed Worsening and Improving subphenotypes were common, had intermediate outcomes across our cohorts, and more nuanced differences in clinical characteristics. These trajectories may be influenced by non-resolving inflammation [30] or immune paralysis [31]. Notably, at 72 hours subjects in the Delayed Worsening subphenotype had diminished leukocyte counts and lymphocyte percentages compared with other groups, which may reflect pathologic T-cell mechanisms described in immune paralysis. These include increased expression of the inhibitory receptor programmed cell death protein 1 (PD-1) and clonal deletion of pathogen-specific cells – mechanisms that may ultimately result in attenuated cell proliferation and immune exhaustion [32]. Further understanding of the biology underlying these subphenotypes will be critical to develop the next generation of treatments for sepsis in all of its forms.

Limitations

This study has several limitations. First, our sepsis subphenotypes were identified based on the data-driven method, which may not be directly related to underlying differences in biology. Biologically derived subphenotypes may help refine our understanding of differential disease progression and the potential for therapeutics to alter the course. Second, because phenotypes were derived in the background of standard of care therapy, it is unclear whether specific processes of care influenced the development of these trajectories. We believe that the stability of these subphenotypes despite secular trends across eleven institutions spanning two decades adds to the generalizability of our findings. Third, although we used eleven separate hospitals in validation, each center is located in the United States and affiliated with an academic center, which may limit generalizability to other locations of care. Lastly, we did not evaluate the effect of specific randomized interventions on SOFA score trajectory.

Conclusions

We have discovered four novel SOFA score trajectory sepsis subphenotypes with different natural histories following admission to the ICU. Our results suggest that these four subphenotypes represent a differential host pathogen response in the setting of current standard of care therapy. Further understanding of the underlying biology of trajectory based subphenotypes may reveal insights into sepsis pathophysiology and improve the personalization of sepsis management and the predictive enrichment of clinical trials.

Table and Figures in Main Text:

Table:

Table 1 Patient characteristics comparisons between subphenotypes

Figures:

Fig. 1 Workflow of study

Fig. 2 Sequential Organ Failure Assessment (SOFA) trajectories of the subphenotypes and survival analysis in terms of the identified subphenotypes. A: Delayed Improving; B: Rapidly Improving; C: Delayed Worsening; D: Rapidly Worsening. The (a) and (b) describe the SOFA trajectories of the subphenotypes on development and validation cohorts, respectively. The (c) and (d) show the survival analysis results in subphenotypes on development and validation cohorts, respectively.

Fig. 3 Chord diagrams showing abnormal variables by subphenotype. A: Delayed Improving; B: Rapidly Improving; C: Delayed Worsening; D: Rapidly Worsening; a: abnormal biomarkers vs. all subphenotypes; I: abnormal biomarkers vs. Delayed Improving; II: abnormal biomarkers vs. Rapidly Improving; III: abnormal biomarkers vs. Delayed Worsening; IV: abnormal biomarkers vs. Rapidly Worsening; b: abnormal subscores vs. all subphenotypes; V: abnormal subscores vs. Delayed Improving; VI: abnormal subscores vs. Rapidly Improving; VII: abnormal subscores vs. Delayed Worsening; VIII: abnormal subscores vs. Rapidly Worsening.

Fig. 4 Predictor contribution to the prediction of the predictive model by using the SHAP technique. Features' importance is ranked based on SHAP values. In this figure, each point represented a single observation. The horizontal location showed whether the effect of that value was associated with a positive (a SHAP value greater than 0) or negative (a SHAP value less than 0) impact on prediction. Color showed whether the original value of that variable was high (in red) or low (in blue) for that observation. For example, in D, a low Platelets value had a positive impact on the Rapidly Worsening subphenotype prediction; the "low" came from the blue color, and the "positive" impact was shown on the horizontal axis. A: Delayed Improving; B: Rapidly Improving; C: Delayed Worsening; D: Rapidly Worsening.

Author contributions

ES, Yuan and FW for conceptualization, investigation, writing, reviewing and editing of the manuscript. ZX for data analysis, drafting, editing and reviewing manuscript. CM for data analysis, editing and reviewing manuscript. CS, IS, LT and DP for discussion, commenting and editing the manuscript.

Data and Code availability

The data and code are available upon request.

Acknowledgements

The work of ZX, CZ and FW are supported by NSF 1750326, NIH RF1AG072449 and NIH R01MH124740. The work of CM and YL are supported in part by NIH 1R01LM013337.

IRB Information

The research is approved by Institutional Review Board of Northwestern University STU00202840.

Declarations

Conflicts of interest

ES reports a conflict for work after January of 2021 and have received personal fees for consulting work for Axle informatics for COVID vaccine clinical trials through NIAID. The other authors declare that they have no conflict of interest.

References:

1. Bhavani SV, Carey KA, Gilbert ER, Afshar M, Verhoef PA, Churpek MM, (2019) Identifying novel sepsis subphenotypes using temperature trajectories. American journal of respiratory and critical care medicine 200: 327-335
2. Rhee C, Dantes R, Epstein L, Murphy DJ, Seymour CW, Iwashyna TJ, Kadri SS, Angus DC, Danner RL, Fiore AE, Jernigan JA, Martin GS, Septimus E, Warren DK, Karcz A, Chan C, Menchaca JT, Wang R, Gruber S, Klompas M, Program ftCPE, (2017) Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009-2014. JAMA 318: 1241-1249
3. Mullard A, (2011) Drug withdrawal sends critical care specialists back to basics. Lancet 378: 1769-1769
4. Seymour CW, Kennedy JN, Wang S, Chang C-CH, Elliott CF, Xu Z, Berry S, Clermont G, Cooper G, Gomez H, Huang DT, Kellum JA, Mi Q, Opal SM, Talisa V, van der Poll T, Visweswaran S, Vodovotz Y, Weiss JC, Yealy DM, Yende S, Angus DC, (2019) Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. JAMA 321: 2003-2017
5. Abraham E, (2021) Moving forward with refinement of definitions for sepsis. Crit Care Med 49: 861-863
6. Cohen J, Opal S, Calandra T, (2012) Sepsis studies need new direction. Lancet Infect Dis 12: 503-505
7. Cohen J, Vincent JL, Adhikari NK, Machado FR, Angus DC, Calandra T, Jaton K, Giulieri S, Delaloye J, Opal S, Tracey K, van der Poll T, Pelfrene E, (2015) Sepsis: a roadmap for future research. Lancet Infect Dis 15: 581-614
8. DeMerle KM, Angus DC, Baillie JK, Brant E, Calfee CS, Carcillo J, Chang CH, Dickson R, Evans I, Gordon AC, Kennedy J, Knight JC, Lindsell CJ, Liu V, Marshall JC, Randolph AG, Scicluna BP, Shankar-Hari M, Shapiro NI, Sweeney TE, Talisa VB, Tang B, Thompson BT, Tsalik EL, van der Poll T, van Vught LA, Wong HR, Yende S, Zhao H, Seymour CW, (2021) Sepsis subclasses: a framework for development and interpretation. Crit Care Med 49: 748-759
9. Knox DB, Lanspa MJ, Kuttler KG, Brewer SC, Brown SM, (2015) Phenotypic clusters within sepsis-associated multiple organ dysfunction syndrome. Intens Care Med 41: 814-822
10. Scicluna BP, van Vught LA, Zwinderman AH, Wiewel MA, Davenport EE, Burnham KL, Nurnberg P, Schultz MJ, Horn J, Cremer OL, Bonten MJ, Hinds CJ, Wong HR, Knight JC, van der Poll T, Consortium M, (2017) Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study. Lancet Resp Med 5: 816-826
11. Sweeney TE, Azad TD, Donato M, Haynes WA, Perumal TM, Henao R, Bermejo-Martin JF, Almansa R, Tamayo E, Howrylak JA, Choi A, Parnell GP, Tang B, Nichols M, Woods CW, Ginsburg GS, Kingsmore SF, Omberg L, Mangravite LM, Wong HR, Tsalik EL, Langley RJ, Khatri P, (2018) Unsupervised analysis of transcriptomics in bacterial sepsis across multiple datasets reveals three robust clusters. Critical Care Medicine 46: 915-925

12. Ferreira FL, Bota DP, Bross A, Melot C, Vincent JL, (2001) Serial evaluation of the SOFA score to predict outcome in critically ill patients. Jama-J Am Med Assoc 286: 1754-1758
13. Vincent JL, Moreno R, Takala J, Willatts S, DeMendonca A, Bruining H, Reinhart CK, Suter PM, Thijs LG, (1996) The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. Intens Care Med 22: 707-710
14. Berndt DJ, Clifford J (1994) Using dynamic time warping to find patterns in time series. In: Editor (ed)^(eds) Book Using dynamic time warping to find patterns in time series. Seattle, WA, USA:, City, pp. 359-370
15. Johnson AEW, Pollard TJ, Shen L, Lehman LWH, Feng ML, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG, (2016) MIMIC-III, a freely accessible critical care database. Sci Data 3
16. Starren JB, Winter AQ, Lloyd-Jones DM, (2015) Enabling a learning health system through a unified enterprise data warehouse: the experience of the northwestern university clinical and translational sciences (NUCATS) institute. Clin Transl Sci 8: 269-271
17. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche J-D, Coopersmith CM, Hotchkiss RS, Levy MM, Marshall JC, Martin GS, Opal SM, Rubenfeld GD, van der Poll T, Vincent J-L, Angus DC, (2016) The third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA 315: 801-810
18. Jain AK, Murty MN, Flynn PJ, (1999) Data clustering: a review. ACM Computing Surveys 31: 264–323
19. McClain JO, Rao VR, (1975) CLUSTISZ: a program to test for the quality of clustering of a set of objects. Journal of Marketing Research 12: 456-460
20. Molnar C (2020) Interpretable machine learning. Lulu. com,
21. Quan HD, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE, Ghali WA, (2005) Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 43: 1130-1139
22. Schenck EJ, Oromendia C, Torres LK, Berlin DA, Choi AMK, Siempos II, (2019) Rapidly improving ARDS in therapeutic randomized controlled trials. Chest 155: 474-482
23. Woźnica EA, Inglot M, Woźnica RK, Łysenko LJAic, University emooWM, (2018) Liver dysfunction in sepsis27: 547-551
24. Gaudry S, Hajage D, Benichou N, Chaibi K, Barbar S, Zarbock A, Lumlertgul N, Wald R, Bagshaw SM, Srisawat N, Combes A, Geri G, Jamale T, Dechartres A, Quenot JP, Dreyfuss D, (2020) Delayed versus early initiation of renal replacement therapy for severe acute kidney injury: a systematic review and individual patient data meta -analysis of randomised clinical trials. Lancet 395: 1506-1515
25. Abrams D, Montesi SB, Moore SKL, Manson DK, Klipper KM, Case MA, Brodie D, Beitler JR, (2020) Powering bias and clinically important treatment effects in randomized trials of critical illness 48: 1710-1719
26. Hotchkiss RS, Moldawer LL, Opal SM, Reinhart K, Turnbull IR, Vincent J-L, (2016) Sepsis and septic shock. Nature Reviews Disease Primers 2: 16045
27. Ma KC, Schenck EJ, Siempos II, Cloonan SM, Finkelzstein EJ, Pabon MA, Oromendia C, Ballman KV, Baron RM, Fredenburgh LEJJi, (2018) Circulating RIPK3 levels are associated with mortality and organ failure during critical illness3
28. Linkermann A, (2019) Death and fire—the concept of necroinflammation. Cell Death & Differentiation 26: 1-3
29. Schenck EJ, Ma KC, Price DR, Nicholson T, Oromendia C, Gentzler ER, Sanchez E, Baron RM, Fredenburgh LE, Huh J-WJJi, (2019) Circulating cell death biomarker TRAIL is associated with increased organ dysfunction in sepsis4
30. Mira JC, Gentile LF, Mathias BJ, Efron PA, Brakenridge SC, Mohr AM, Moore FA, Moldawer LL, (2017) Sepsis pathophysiology, chronic critical illness, and persistent inflammation-immunosuppression and catabolism syndrome45: 253-262
31. Cao C, Yu M, Chai Y, (2019) Pathological alteration and therapeutic implications of sepsis-induced immune cell apoptosis. Cell Death & Disease 10: 782

32. Yi JS, Cox MA, Zajac AJ, (2010) T-cell exhaustion: characteristics, causes and conversion. Immunology 129: 474-481

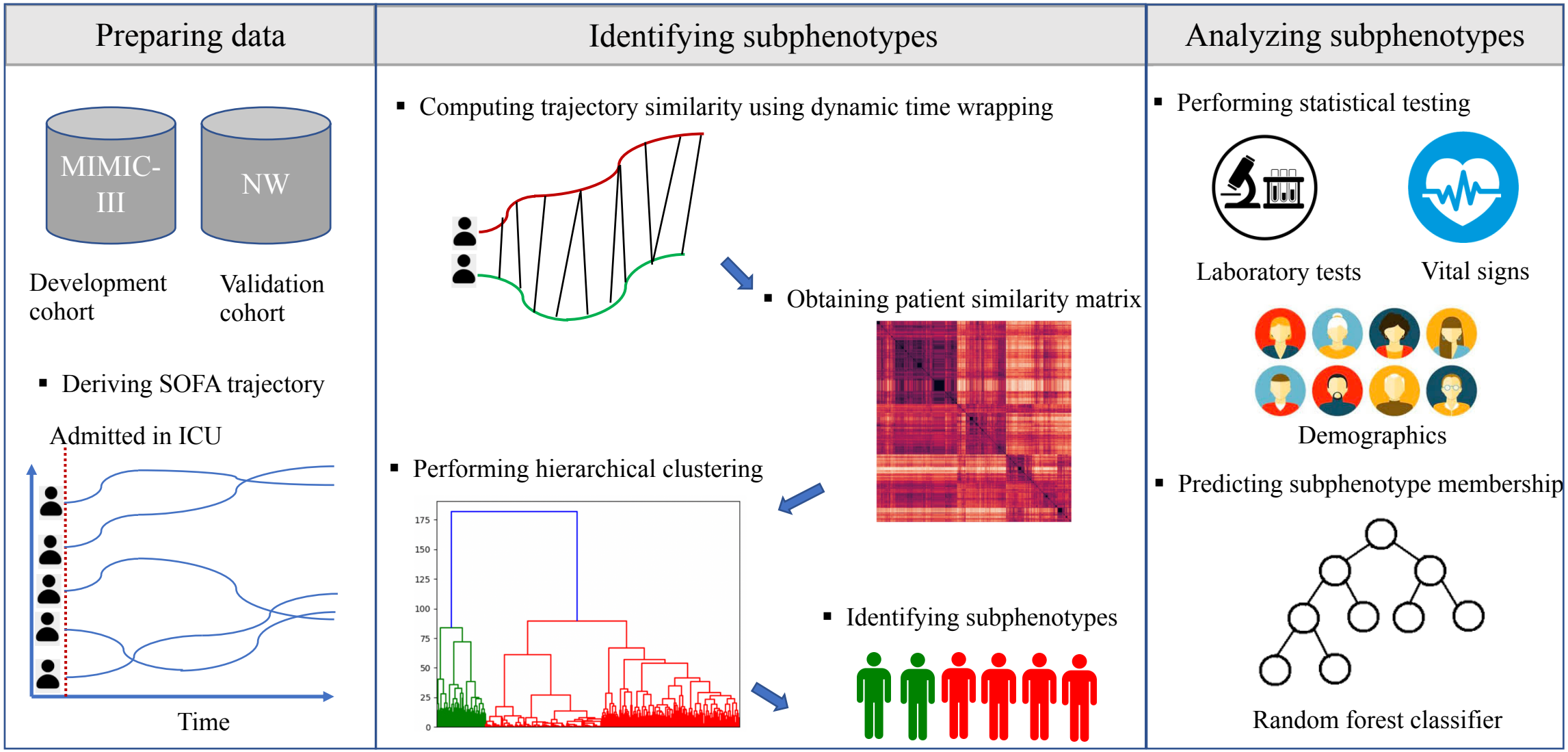
Preparing data
MIMIC-III
NW
Development cohort
Validation cohort
Deriving SOFA trajectory
Admitted in ICU
Time
Identifying subphenotypes
Computing trajectory similarity using dynamic time wrapping
Obtaining patient similarity matrix
Performing hierarchical clustering
175
150
125
100
75
50
25
0
Identifying subphenotypes
Analyzing subphenotypes
Performing statistical testing
Laboratory tests
Vital signs
Demographics
Predicting subphenotype membership
Random forest classifier

Fig. 1 Workflow of study

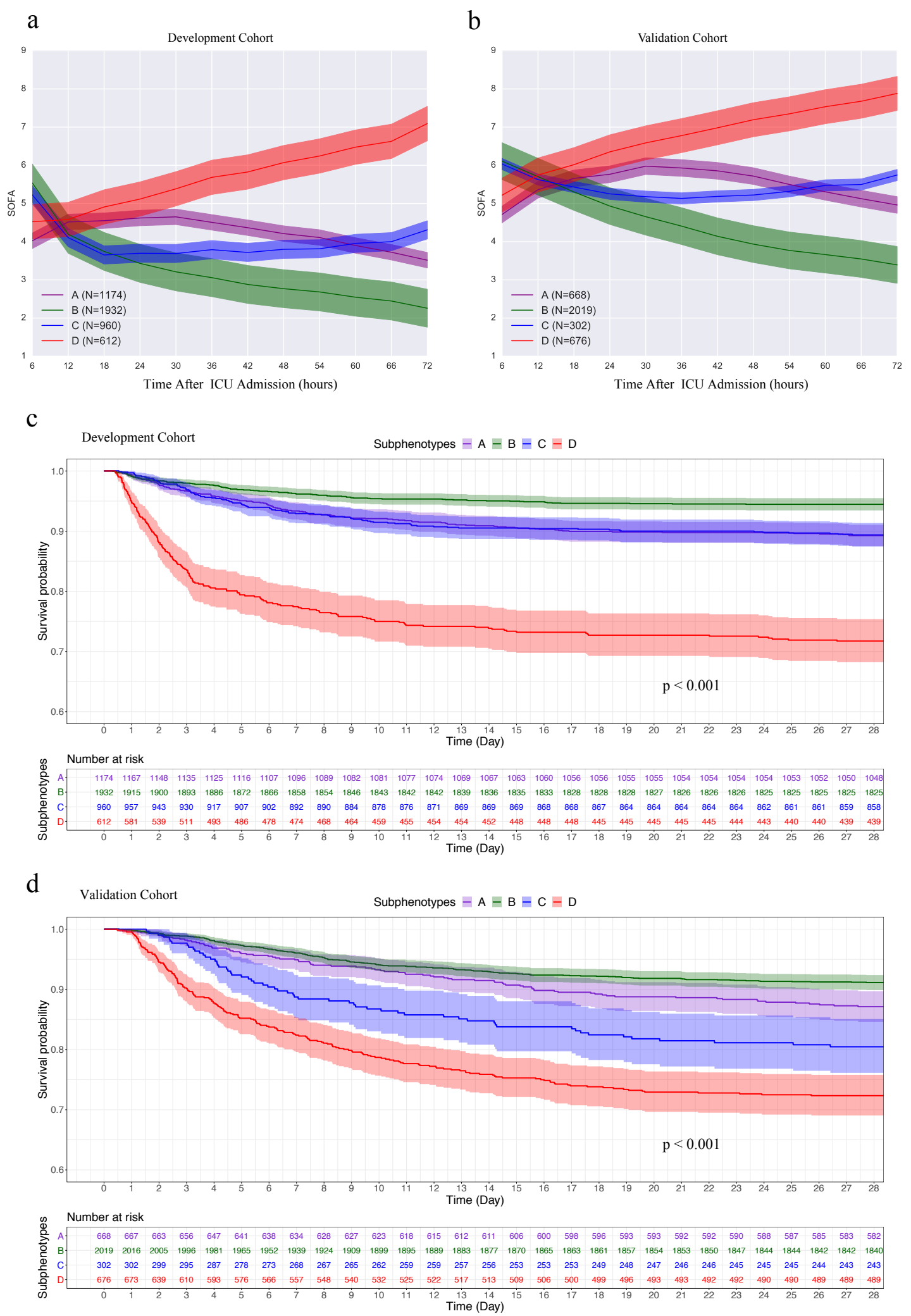
a
Development Cohort
SOFA
A (N=1174)
B (N=1932)
C (N=960)
D (N=612)
Time After ICU Admission (hours)
b
Validation Cohort
SOFA
A (N=668)
B (N=2019)
C (N=302)
D (N=676)
Time After ICU Admission (hours)
c
Development Cohort
Subphenotypes A B C D
Survival probability
p < 0.001
Time (Day)
Number at risk
Subphenotypes
A 1174 1167 1148 1135 1125 1116 1107 1096 1089 1082 1081 1077 1074 1069 1067 1063 1060 1056 1056 1055 1055 1054 1054 1054 1054 1053 1052 1050 1048
B 1932 1915 1900 1893 1886 1872 1866 1858 1854 1846 1843 1842 1842 1839 1836 1835 1833 1828 1828 1828 1827 1826 1826 1826 1825 1825 1825 1825 1825
C 960 957 943 930 917 907 902 892 890 884 878 876 871 869 869 869 868 868 867 864 864 864 864 864 862 861 861 859 858
D 612 581 539 511 493 486 478 474 468 464 459 455 454 454 452 448 448 448 445 445 445 445 445 444 443 440 440 439 439
Time (Day)
d
Validation Cohort
Subphenotypes A B C D
Survival probability
p < 0.001
Time (Day)
Number at risk
Subphenotypes
A 668 667 663 656 647 641 638 634 628 627 623 618 615 612 611 606 600 598 596 593 593 592 592 590 588 587 585 583 582
B 2019 2016 2005 1996 1981 1965 1952 1939 1924 1909 1899 1895 1889 1883 1877 1870 1865 1863 1861 1857 1854 1853 1850 1847 1844 1844 1842 1842 1840
C 302 302 299 295 287 278 273 268 267 265 262 259 259 257 256 253 253 253 249 248 247 246 246 245 245 245 244 243 243
D 676 673 639 610 593 576 566 557 548 540 532 525 522 517 513 509 506 500 499 496 493 493 492 492 490 490 489 489 489
Time (Day)

Fig. 2 Sequential Organ Failure Assessment (SOFA) trajectories of the subphenotypes and survival analysis in terms of the identified subphenotypes. A: Delayed Improving; B: Rapidly Improving; C: Delayed Worsening; D: Rapidly Worsening. The (a) and (b) describe the SOFA trajectories of the subphenotypes on development and validation cohorts, respectively. The (c) and (d) show the survival analysis results in subphenotypes on development and validation cohorts, respectively.

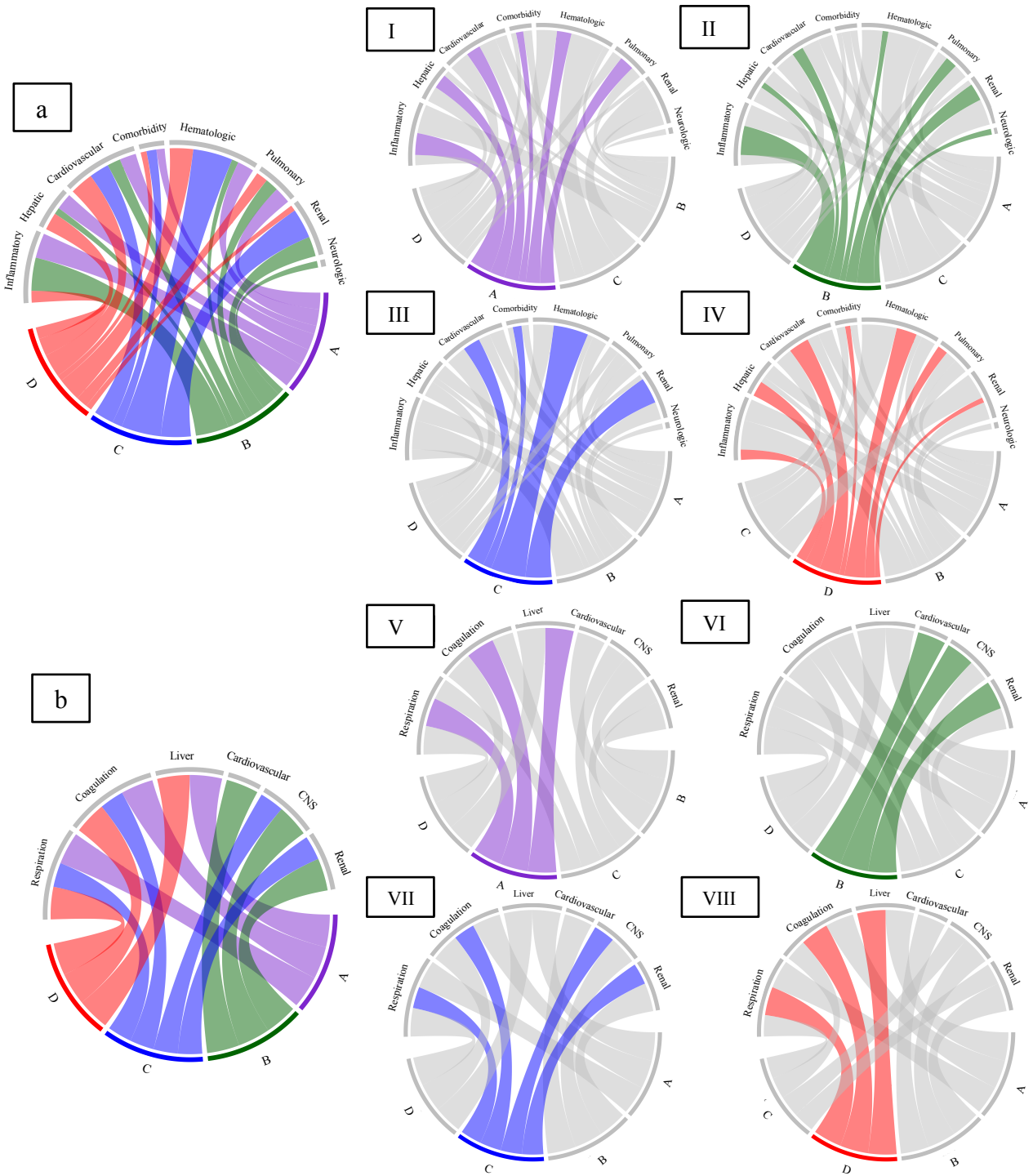
a
Inflammatory
Hepatic
Cardiovascular
Comorbidity
Hematologic
Pulmonary
Renal
Neurologic
A
B
C
D
I
II
III
IV
b
Respiration
Coagulation
Liver
Cardiovascular
CNS
Renal
V
VI
VII
VIII

Fig. 3 Chord diagrams showing abnormal variables by subphenotype. A: Delayed Improving; B: Rapidly Improving; C: Delayed Worsening; D: Rapidly Worsening; a: abnormal biomarkers vs. all subphenotypes; I: abnormal biomarkers vs. Delayed Improving; II: abnormal biomarkers vs. Rapidly Improving; III: abnormal biomarkers vs. Delayed Worsening; IV: abnormal biomarkers vs. Rapidly Worsening; b: abnormal subscores vs. all subphenotypes; V: abnormal subscores vs. Delayed Improving; VI: abnormal subscores vs. Rapidly Improving; VII: abnormal subscores vs. Delayed Worsening; VIII: abnormal subscores vs. Rapidly Worsening.

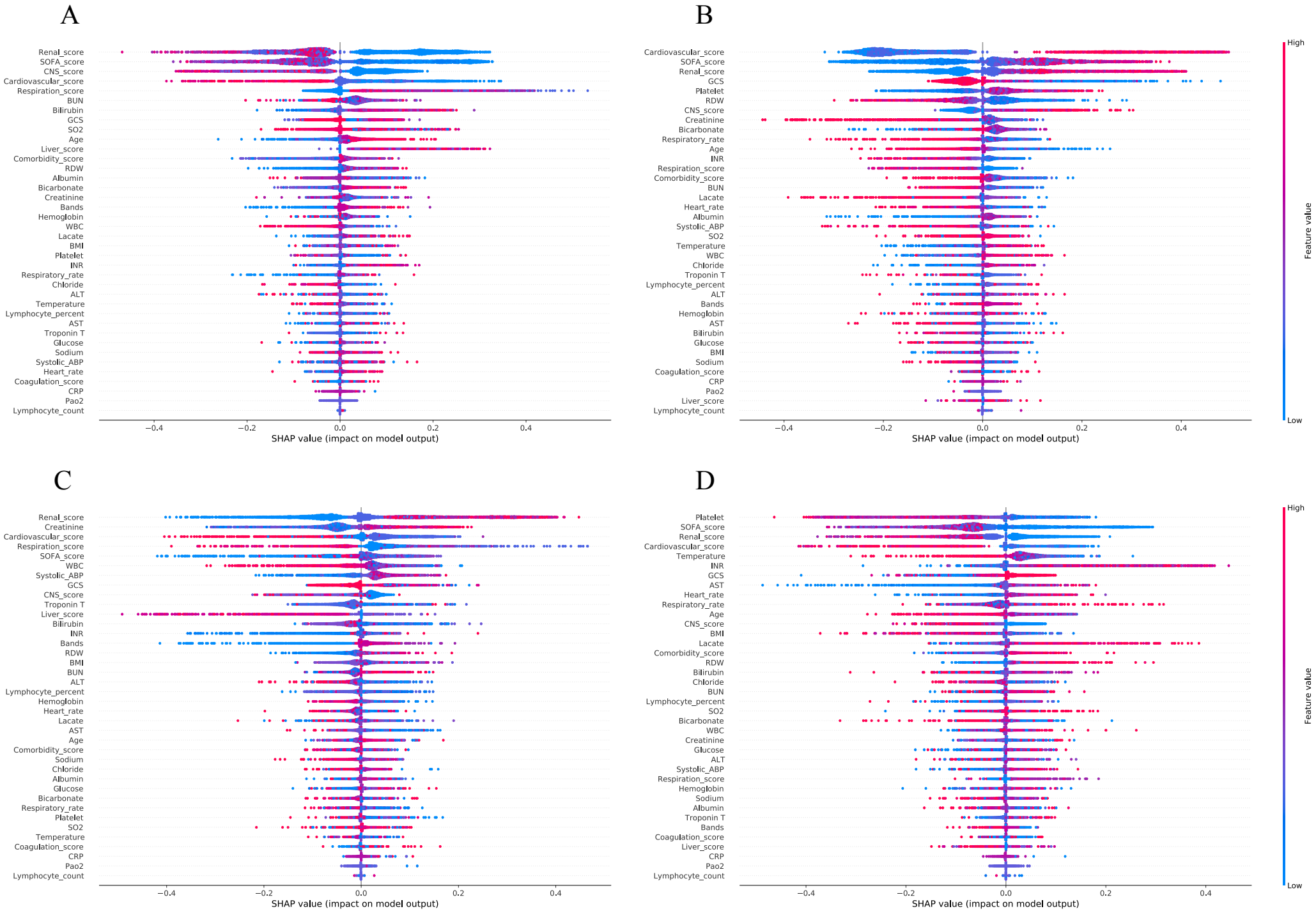
A
Renal_score
SOFA_score
CNS_score
Cardiovascular_score
Respiration_score
BUN
Bilirubin
GCS
SO2
Age
Liver_score
Comorbidity_score
RDW
Albumin
Bicarbonate
Creatinine
Bands
Hemoglobin
WBC
Lacate
BMI
Platelet
INR
Respiratory_rate
Chloride
ALT
Temperature
Lymphocyte_percent
AST
Troponin T
Glucose
Sodium
Systolic_ABP
Heart_rate
Coagulation_score
CRP
Pao2
Lymphocyte_count
−0.4
−0.2
0.0
0.2
0.4
SHAP value (impact on model output)
B
Cardiovascular_score
SOFA_score
Renal_score
GCS
Platelet
RDW
CNS_score
Creatinine
Bicarbonate
Respiratory_rate
Age
INR
Respiration_score
Comorbidity_score
BUN
Lacate
Heart_rate
Albumin
Systolic_ABP
SO2
Temperature
WBC
Chloride
Troponin T
Lymphocyte_percent
ALT
Bands
Hemoglobin
AST
Bilirubin
Glucose
BMI
Sodium
Coagulation_score
CRP
Pao2
Liver_score
Lymphocyte_count
−0.4
−0.2
0.0
0.2
0.4
SHAP value (impact on model output)
High
Feature value
Low
C
Renal_score
Creatinine
Cardiovascular_score
Respiration_score
SOFA_score
WBC
Systolic_ABP
GCS
CNS_score
Troponin T
Liver_score
Bilirubin
INR
Bands
RDW
BMI
BUN
ALT
Lymphocyte_percent
Hemoglobin
Heart_rate
Lacate
AST
Age
Comorbidity_score
Sodium
Chloride
Albumin
Glucose
Bicarbonate
Respiratory_rate
Platelet
SO2
Temperature
Coagulation_score
CRP
Pao2
Lymphocyte_count
−0.4
−0.2
0.0
0.2
0.4
SHAP value (impact on model output)
D
Platelet
SOFA_score
Renal_score
Cardiovascular_score
Temperature
INR
GCS
AST
Heart_rate
Respiratory_rate
Age
CNS_score
BMI
Lacate
Comorbidity_score
RDW
Bilirubin
Chloride
BUN
Lymphocyte_percent
SO2
Bicarbonate
WBC
Creatinine
Glucose
ALT
Systolic_ABP
Respiration_score
Hemoglobin
Sodium
Albumin
Troponin T
Bands
Coagulation_score
Liver_score
CRP
Pao2
Lymphocyte_count
−0.4
−0.2
0.0
0.2
0.4
SHAP value (impact on model output)
High
Feature value
Low

Fig. 4 Predictor contribution to the prediction of the predictive model by using the SHAP technique. Features' importance is ranked based on SHAP values. In this figure, each point represented a single observation. The horizontal location showed whether the effect of that value was associated with a positive (a SHAP value greater than 0) or negative (a SHAP value less than 0) impact on prediction. Color showed whether the original value of that variable was high (in red) or low (in blue) for that observation. For example, in D, a low Platelets value had a positive impact on the Rapidly Worsening subphenotype prediction; the "low" came from the blue color, and the "positive" impact was shown on the horizontal axis. A: Delayed Improving; B: Rapidly Improving; C: Delayed Worsening; D: Rapidly Worsening.

Table 1 Patient characteristics comparisons between subphenotypes

Characteristics	Total N=4,678	Delayed Improving N=1,174	Rapidly Improving N=1,932	Delayed Worsening N=960	Rapidly Worsening N=612	p-value†
Age, median (IQR)	65.93 [53.71-77.88]	67.25[54.84-79.17]	65.33 [53.26-77.2]	66.95 [53.94-78.32]	64.52 [52.52-76.68]	0.204
Gender, No. (%)						
Male	2625 (56.11)	594 (50.60)	1100 (56.94)	548 (57.08)	383 (62.58)	0.081
Female	2053 (43.89)	580 (49.40)	832 (43.06)	412 (42.92)	229 (37.42)	
Race, No. (%)						0.207
WHITE	3367 (71.98)	870 (74.11)	1398 (72.36)	670 (69.79)	429 (70.10)	
BLACK	424 (9.06)	92 (7.84)	189 (9.78)	101(10.52)	42 (6.86)	
OTHER	887 (18.96)	212 (18.06)	345 (17.86)	189 (19.69)	141(23.04)	
Elixhauser index, median (IQR)	4.0 [0.0-9.0]	4.0 [0.0-9.0]	4.0 [0.0-9.0]	4.0 [0.0-9.0]	5.0 [0.0-10.0]	0.015
Length stay, median (IQR)	2.76 [1.59-5.57]	2.97 [1.76-6.15]	2.42 [1.46-4.78]	2.88 [1.65-5.29]	2.88 [1.57-6.7]	< 0.001
Mechanical ventilation at admission, No. (%)	1893 (40.47)	499 (42.50)	733 (37.94)	377 (39.27)	284 (46.41)	< 0.001
Baseline SOFA, mean (SD)	4.96 (2.80)	4.02 (2.36)	5.54 (2.92)	5.23 (2.73)	4.52 (2.80)	< 0.001
ICU unit at admission, No. (%)						0.037
SICU	771 (16.48)	185 (15.76)	341 (17.65)	135 (14.06)	110 (17.97)	
CCU	443 (9.47)	117 (9.97)	167 (8.64)	94 (9.79)	65 (10.62)	
TSICU	593 (12.68)	173 (14.74)	226 (11.70)	119 (12.40)	75 (12.25)	
MICU	2611 (55.81)	634 (54.00)	1087 (56.26)	569 (59.27)	321 (52.45)	
CSRU	260 (5.56)	65 (5.54)	111 (5.75)	43 (4.48)	41 (6.70)	
Admission location, No. (%)						0.196
Transfer from HOSP/EXTRAM	810 (17.32)	213 (18.14)	304 (15.73)	165 (17.19)	128 (20.92)	
Emergency room ADMIT	1497 (32.00)	355 (30.24)	628 (32.51)	328 (34.17)	186 (30.39)	
Clinic referral/premature	1985 (42.43)	493 (41.99)	847 (43.84)	396 (41.25)	249 (40.69)	
Transfer from other HEALT	4 (0.09)	2 (0.17)	1 (0.05)	1 (0.10)	0 (0.00)	
PHYS referral/normal DELI	367 (7.85)	106 (9.03)	145 (7.51)	69 (7.19)	47 (7.68)	
Transfer from skilled NUR	15 (0.32)	5 (0.43)	7 (0.36)	1 (0.10)	2 (0.33)	
Infection item, No. (%)						
cfi_Central nervous system	56 (1.20)	10 (0.85)	27 (1.40)	8 (0.83)	11 (1.80)	0.189
cfi_Intra-Abdominal	880 (18.81)	230 (19.59)	363 (18.79)	172 (17.92)	115 (18.79)	0.808
cfi_Pneumonia	1257 (26.87)	328 (27.94)	494 (25.57)	262 (27.29)	173 (28.27)	0.385
cfi_Septicemia bacteremia	1587 (33.92)	359 (30.58)	717 (37.11)	300 (31.25)	211 (34.48)	< 0.001
cfi_Skin soft tissue	276 (5.90)	60 (5.11)	140 (7.25)	42 (4.38)	34 (5.56)	0.008
cfi_Urinary Tract	1044 (22.32)	276 (23.51)	439 (22.72)	228 (23.75)	101 (16.50)	0.003
Septic shock, No. (%)	635(13.57%)	148(12.61%)	299(15.48%)	101(10.52%)	87(14.22%)	0.002

Definition of abbreviations: IQR--interquartile range; SD--standard deviation; SOFA--Sequential Organ Failure Assessment; SICU-- Surgical ICU; CCU-- Coronary Care Unit; TSICU--Thoracic Surgery ICU; MICU--Medical ICU; CSRU--Cardiac Surgery ICU. †p-value calculated by Chi-square test/Fisher's exact test, or student's t-test/Mann-Whitney test where appropriate.