Subphenotyping heterogeneous patients with chronic critical illness to guide individualised fluid balance treatment using machine learning: a retrospective cohort study

Peizhao Liu,[a,e] Sicheng Li,[a,e] Tao Zheng,[c,e] Jie Wu,[c] Yong Fan,[d] Xiaoli Liu,[d] Wenbin Gong,[b] Haohao Xie,[c] Juanhan Liu,[a] Yangguang Li,[a] Haiyang Jiang,[c] Fan Zhao,[a] Jinpeng Zhang,[a] Lei Wu,[a] Huajian Ren,[a] Zhiwu Hong,[a] Jun Chen,[a] Guosheng Gu,[a] Gefei Wang,[a] Zhengbo Zhang,[d,*] Xiuwen Wu,[a,**] Yun Zhao,[c,***] and Jianan Ren[a,****]

[a]Research Institute of General Surgery, Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, China
[b]School of Medicine, Southeast University, Nanjing, 210002, China
[c]Department of General Surgery, BenQ Medical Center, The Affiliated BenQ Hospital of Nanjing Medical University, Nanjing, 210019, China
[d]Center for Artificial Intelligence in Medicine, Chinese PLA General Hospital, Beijing, 100039, China

Summary

Background The great heterogeneity of patients with chronic critical illness (CCI) leads to difficulty for intensive care unit (ICU) management. Identifying subphenotypes could assist in individualized care, which has not yet been explored. In this study, we aim to identify the subphenotypes of patients with CCI and reveal the heterogeneous treatment effect of fluid balance for them.

Methods In this retrospective study, we defined CCI as an ICU length of stay over 14 days and coexists with persistent organ dysfunction (cardiovascular Sequential Organ Failure Assessment (SOFA) score ≥1 or score in any other organ system ≥2) at Day 14. Data from five electronic healthcare record datasets covering geographically distinct populations (the US, Europe, and China) were studied. These five datasets include (1) subset of Derivation (MIMIC-IV v1.0, US) cohort (2008–2019); (2) subset Derivation (MIMIC-III v1.4 'CareVue', US) cohort (2001–2008); (3) Validation I (eICU-CRD, US) cohort (2014–2015); (4) Validation II (AmsterdamUMCdb/AUMC, Euro) cohort (2003–2016); (5) Validation III (Jinling, CN) cohort (2017–2021). Patients who meet the criteria of CCI in their first ICU admission period were included in this study. Patients with age over 89 or under 18 years old were excluded. Three unsupervised clustering algorithms were employed independently for phenotypes derivation and validation. Extreme Gradient Boosting (XGBoost) was used for phenotype classifier construction. A parametric G-formula model was applied to estimate the cumulative risk under different daily fluid management strategies in different subphenotypes of ICU mortality.

Findings We identified four subphenotypes as Phenotype A, B, C, and D in a total of 8145 patients from three countries. Phenotype A is the mildest and youngest subgroup; Phenotype B is the most common group, of whom patients showed the oldest age, significant acid-base abnormality, and low white blood cell count; Patients with Phenotype C have hypernatremia, hyperchloremia, and hypercatabolic status; and in Phenotype D, patients accompany with the most severe multiple organ failure. An easy-to-use classifier showed good effectiveness. Phenotype characteristics showed robustness across all cohorts. The beneficial fluid balance threshold intervals of subphenotypes were different.

Interpretation We identified four novel phenotypes that revealed the different patterns and significant heterogeneous treatment effects of fluid therapy within patients with CCI. A prospective study is needed to validate our findings, which could inform clinical practice and guide future research on individualized care.

*Corresponding author. Centre for Artificial Intelligence in Medicine, Chinese PLA General Hospital, Beijing, China.
**Corresponding author. Research Institute of General Surgery, Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, China.
***Corresponding author. Department of General Surgery, BenQ Medical Centre, The Affiliated BenQ Hospital of Nanjing Medical University, Nanjing, China.
****Corresponding author. Research Institute of General Surgery, Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, China.
    E-mail addresses: zhengbozhang@126.com (Z. Zhang), wuxiuwen@nju.edu.cn (X. Wu), zhaoyun056@gmail.com (Y. Zhao), jiananr@nju.edu.cn (J. Ren).
[e]Peizhao Liu, Sicheng Li, Tao Zheng contributed equally to this work.
Translation For the Chinese translation of the Summary, see the Supplementary Materials section.

Research in context

Evidence before this study

The great heterogeneity of patients with chronic critical illness (CCI) makes it very difficult to implement appropriate diagnosis and fluid treatment. We firstly searched PubMed, Google Scholar, and preprint platforms (medRxiv, bioRxiv, arXiv, and Research Square) by using the term "[("chronic critical illness" OR "persistent critical illness") AND ("subphenotype" OR "phenotype" OR "subclass" OR "subtype")]" and "["critical care" AND ("fluid therapy" OR "fluid management" OR "fluid balance")]" with no language restrictions between Jan 1, 1990 to Nov 30, 2022. The identified studies have defined a subgroup of patients with CCI as the persistent inflammation immunosuppression and catabolism syndrome (PIICS), and the current findings support that restricted fluid administration for critically ill patient may be beneficial while fluid overload may be harmful. However, these studies showed limited sample size and geographically diversity, significant clinical heterogeneity from each other, and no specific report focus on the fluid management in CCI population.

Added value of this study

Our study is the first that discovered the subphenotypes and path the way to guide individualized treatment in patients with CCI. We used multiple machine learning algorithms by analysed electronic health records and identified four phenotypes in 8145 patients with CCI. The four subphenotypes showed distinct characteristics and outcomes from each other and the overall population. Our study covered a much larger scale of sample size and wide range than previous CCI studies. The reliability and reproducibility of our findings were strongly validated methodologically and geographically (US, China, and Europe). We found these four CCI phenotypes cannot be explained by traditional illness severity measurement such as Sequential Organ Failure Assessment (SOFA) score. We constructed a parsimonious eXtreme Gradient Boosting based phenotype classifier to directly facilitate clinical practice by providing clinicians a user-friendly application. This classifier showed good subphenotype assignment prediction effectness. Not only that, we used a parametric G-formula framework and found the heterogeneous treatment effect of daily fluid management in different CCI subphenotypes in intensive care unit.

Implications of all the available evidence

Our work provides a solution to improve management in patients with CCI that has never been reported before. Based on our findings, the recognition of patients diagnosed with CCI should be more specific, and appropriate fluid management strategies based on different phenotype could improve outcomes.

Introduction

Chronic critical illness (CCI) is an "epidemic" in intensive care unit (ICU), an inevitable consequence of modern medical developments.[1,2] In the past 30 years, the prevalence of CCI has been increasing in the world, especially in the high-income countries, which has attracted academic attention.[3] In US, the overall national population-based prevalence of CCI is 34.4 per 100,000 with increasing growth rate of 25.76% from 2004 to 2009.[4] In Japan, a nationwide inpatients survey showed the overall age-specific population prevalence of CCI was 42.0 per 100,000, with 28.6% inhospital mortality.[5]

The widespread and increasing occurrence of CCI not only endangers patients' health but also causes a heavy burden to their families, the healthcare system, and society due to the high occupancy of valuable resources in ICU.[4–6] Nevertheless, with such enormous burdens and expenditures, the prognoses of patients with CCI remain poor. The in-hospital mortality is approximately 30%,[4,5] and 5-year mortality are around 81%,[7] which is even higher than most malignant cancers (20–49%).[2]

Gardner et al. proposed a criterion of CCI diagnosis as an ICU length of stay over 14 days and coexists with persistent organ dysfunction (cardiovascular Sequential Organ Failure Assessment (SOFA) score ≥1 or score in any other organ system ≥2) at Day 14. Although the awareness and understanding of CCI have advanced considerably, it has not translated into specific therapeutic approaches. Current treatment strategies in CCI are symptomatic and empirical. One of the main reasons hindering the treatment progress of CCI is the existing great heterogeneous pathophysiology.[1,8]

In clinical practice, two patients diagnosed with CCI may hide completely different pathophysiological

states,[1,9] affected by various factors such as age, acid-base balance, body metabolism, inflammatory status, immune response, and degree of systemic damage. Mira et al. proposed the concept of persistent inflammation, immunosuppression, and catabolism syndrome (PIICS) to describe a typical type of CCI patient with initial sepsis and trauma.[10] However, in addition to patients with CCI characterized by PIICS, other phenotypes or subgroups of patients have not yet been studied. These unidentified subgroups or phenotypes may have a different risk of adverse outcomes and respond differently to treatment.[1]

Fluid management in critically ill patients is a topic that has seen a sharp increase in academic attention and concern in recent years. Fluid management during the ICU period has been reported to affect long-term organ function and outcome.[11] Current clinical studies have majorly focused on fluid resuscitation strategies in the acute critical care phase, but little attention has been paid to fluid therapy in the chronic critical care phase. Due to the great heterogeneity of patients with CCI, developing a uniform fluid therapy protocol seems impractical. Therefore, it is highly necessary to explore the heterogeneity among different fluid management strategies to guide individualized management in patients with CCI.

In this study, we first sought to identify and validate patterns among patients with CCI from retrospectively collected electronic healthcare records (EHR) covering different countries (US, Netherlands, and China) using unsupervised machine learning methods. To facilitate clinical use, we constructed a parsimonious classifier for subphenotypes in a user-friendly application based on easy-obtained indicators during daily care. Finally, we sought to explore the heterogeneous correlation between daily fluid balance and prognosis among different CCI subphenotypes. Overall, we aim to identify subphenotypes in patients with CCI and highlight the promising clinical applied prospect in CCI management.

Part of these findings has been previously reported in a conference abstract.[12]

Methods
This study followed the reporting guidelines of the STROBE guidelines. The overall workflow chart was illustrated in Fig. 1.

Dataset and study cohort
We used the EHR data obtained from four public critical care databases and proprietary real-world EHR dataset from Nanjing, China – (1) subset of Derivation (Medical Information Mart for Intensive Care IV/MIMIC-IV v1.0, US) cohort (2008–2019)[13]; (2) subset Derivation (Medical Information Mart for Intensive Care III/MIMIC-III v1.4 'CareVue', US) cohort (2001–2008)[14]; (3) Validation I (eICU Collaborative Research Database/eICU-CRD, US) cohort (2014–2015)[15]; (4) Validation II (Amsterdam University Medical Centre Database/AmsterdamUMCdb/

AUMC, Euro) cohort (2003–2016)[16]; (5) Validation III (Jinling, China) cohort (2017–2021). The Treatment cohort was longitudinal data obtained from MIMIC-IV and eICU-CRD database. The detailed information was shown in Supplementary Methods.

All patients diagnosed with CCI in ICU were included in this study. CCI was defined as a long-term status with critical organ dysfunction during ICU stay, and the criteria we used were: (a) ICU duration ≥14 days; (b) coexist with the evidence of persistent organ dysfunction (cardiovascular SOFA ≥1 or score in any other organ system ≥2) at Day 14.[3,6,17] The exclusion criteria were: (a) Age >89 years old; (b) Age <18 years old; (c) Sequence of ICU admission ≥2. See Supplementary Fig. S1. The SOFA scores were calculated by the worst value of variables measured during Day 14.

### Candidate clinical variables for phenotyping
Variables were extracted from datasets on Day 14 during ICU stay. In total 51 clinical variables were obtained from raw datasets. After evaluating the missing value proportion (Supplementary Fig. S2), 25 variables remained for further analysis after removing missingness over 40% variables. These variables include demographic information (age/age groups), vital signs (mean blood pressure, heart rate, respiratory rate, temperature, Glasgow coma scale, oxygen saturation, fractional inspired oxygen, urine output), laboratory indicators (anion gap, bicarbonate, chloride, sodium, potassium, calcium, magnesium, blood urea nitrogen, creatinine, platelet, haemoglobin, red cell distribution width, mean corpuscular volume, urea-creatinine ratio, white blood cell, glucose). We also calculated the urea-creatinine ratio as a potential catabolic indicator.[18] For other indicators with multiple records, mean values were calculated and analysed. We categorized all variables into corresponding organ systems and defined the abnormality direction of each variable (Supplementary Table S1).

Observational endpoints
The primary endpoint was death during ICU stay. Secondary endpoints included in-hospital death, ICU and hospital length of stay, and hospital discharge location. We unified the terms of discharge location information of different datasets and divided these destinations into two categories – the good disposition and the poor disposition (Supplementary Table S2).[19] For the AmsterdamUMCdb dataset, only ICU mortality and ICU length of stay were analysed.

Derivation and validation of subphenotypes
Before clustering analysis, we performed data pre-processing procedures. There are some severe deviations, errors, or missing values in our cohorts. First, we replaced those outliers and error values with
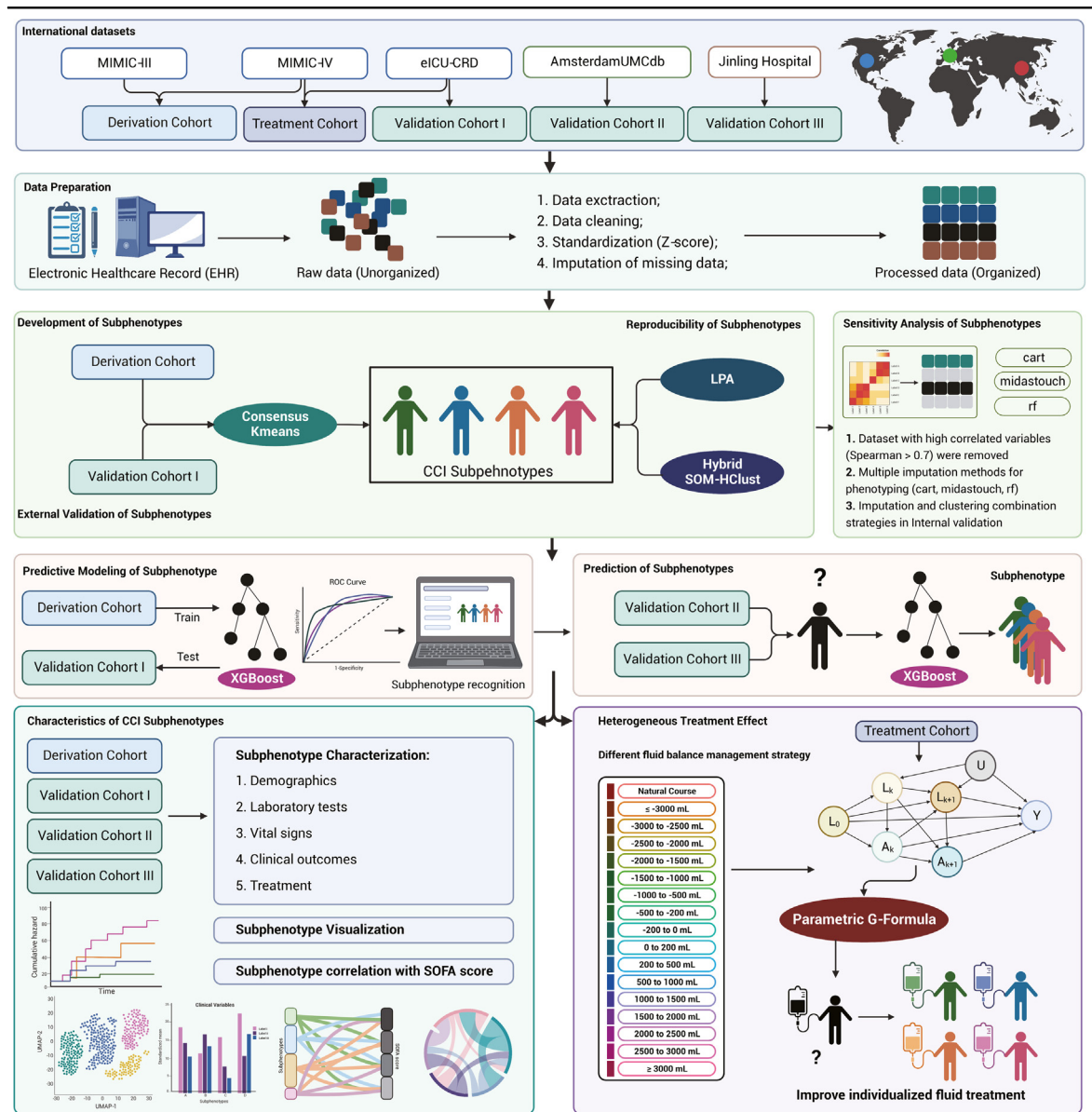
Fig. 1: Schematic of the study workflow. The schematic of the workflow of the study. First, data obtained from multiple resources capable of representing the different countries were extracted. Several data preparation strategies were employed. The datasets were classified as the Derivation cohort, the Validation I cohort, the Validation II cohort, and the Validation III cohort. Consensus Kmeans phenotyping was initially performed on the Derivation cohort and then validated in the Validation I cohort. For reproducibility, latent profile analysis and hybrid SOM-hierarchical clustering phenotyping were used for comparison with consensus Kmeans. UMAP and other visualization protocols were employed for comparisons. Sensitivity analysis included removing highly correlated variables for clustering, phenotyping in data imputed by different methods, and different combination strategies of imputation and clustering in the Derivation cohort. Further analyses were conducted for phenotypes interpretation, the correlation between subphenotypes and SOFA score (to ensure the phenotypes were not simply recapitulations of classical clinical groups and severities). Next, a parsimonious classifier was built for subphenotypes classification, and the model further conducted phenotypes assignment in the Validation II and III cohort. Finally, we analysed the longitudinal data from the Treatment cohort. In directed acyclic graph, the arrow direction defined the potential causal framework for interventions (A_k, A_{k+1}, ...), baseline variables (L_0), time dependent variables (L_k, L_{k+1}, ...), unmeasured covariates (U), and outcome variables (Y). Multiple simulation processes under different treatment strategies of daily fluid balance were analysed to explore the heterogeneous association with ICU survival by using a parametric G-formula compared to the natural course for the overall population and each four phenotypes from the Treatment cohort independently. CCI – Chronic critical illness; PIICS – Persistent inflammation, immunosuppression, and catabolic syndrome; SOFA – Sequential organ failure assessment; SOM – Self-organizing map; HClust – Hierarchical clustering; MIMIC-IV – Medical Information Mart for Intensive Care IV; MIMIC-III –

missing values. Second, variables with high missingness (>40% data missed) from the Derivation and Validation I cohort were removed. FiO2 was imputed by 21%. For the rest of the variables, we employed multiple imputations by using predictive mean matching (pmm), with three additional methods – classification and regression tree (cart), weighted predictive mean matching (midastouch), and random forest imputations (rf) for sensitivity analysis. In total 20 imputed datasets and two merged datasets by computing the mean and median value of each variable were generated. We used the integrated imputed dataset of mean value for subsequential explorative analysis. For more details, please see Supplementary Methods.

To identify CCI subclasses, we firstly assessed and pre-processed the datasets (Supplementary Methods, Figs. S3–S7, S18–S21, and S26). We applied the consensus clustering in the Derivation cohort. The optimal number of clusters was determined under multiple indicators consideration, including (1) the concentration degree of consensus matrix, (2) the flatness of consensus cumulative density function (CDF) curves, (3) the elbow point of the area under the CDF (delta values) curve, and (4) clusters with higher cluster consensus (≥0.8) for all clusters would be considered appropriate. The same phenotyping framework was performed in the Validation I cohort for external validation.

To evaluate the reproducibility of CCI subphenotypes, we performed two different clustering methods – the latent profile analysis (LPA) and the hybrid self-organizing map (SOM)-hierarchical clustering (hSOM). For LPA, the optimal number of clusters was determined by the Bayesian information criterion (BIC), entropy, the bootstrap likelihood ratio test (BLRT), and Lo-Mendell–Rubin test (LMR). We used statistical tests including BLRT-p and LMR-p to determine the optimal number of clusters. If statistical tests were not applicable, we calculated the elbow point of BIC to determine the optimal profiles number. Entropy was used to evaluate model accuracy and a cluster number with entropy close to 0.8 was deemed accurate. We also considered model stability by setting a sample size threshold at 5% of the population's posterior mode for each profile. For hSOM, a SOM object was computed first then we performed a hierarchical clustering, the optimal clusters number was determined by the elbow point of the sum of square changes. More details can be found in the Supplementary Methods.

Subphenotype classifier
We used the eXtreme Gradient Boosting (XGBoost) to model a subphenotype-identification classifier. Variables were selected carefully based on the feature importance to build the final compact classifier. The Derivation cohort was employed as the training set. We used 10-folds cross-validation in the training set for internal validation. The Validation I cohort was used as the testing set for external validation. The Validation II and III cohort phenotypes were evaluated using this model. We plotted multi-class receiver operating characteristic (ROC) curves and calculated the area under the curve (AUC) to evaluate our model performance. The classifier was further packaged into an R Shiny-based application for utility (Supplementary Methods).

Heterogeneous treatment effect of daily fluid balance
To further investigate the potential heterogeneity of fluid therapy in different phenotypes, we extracted and pre-processed the records from CCI diagnosed to ICU discharge and calculated the daily fluid net input in the Treatment cohort. The parametric G-formula model was employed to investigate the independent association between daily fluid balance and ICU survival after CCI diagnosed.[20] The variables included in the model were daily fluid net amount, demographic information (age, sex, ethnicity), daily SOFA score, Charlson comorbidity index, treatments (daily use of albumin infusion, daily use of CRRT, daily diuretics usage and loop diuretics dose, daily use of packed RBC transfusion, daily vasopressor dose, daily use of invasive ventilation), daily vital signs and daily laboratory variables (anion gap, bicarbonate, BUN, chloride, creatinine, haemoglobin, glucose, potassium, sodium, WBC, heart rate, respiratory rate, mean blood pressure, oxygen saturation, temperature, Glasgow coma scale, urea-creatinine ratio). The calculation of daily net fluid intake included all available records in that day. We set a series of different scaled fluid management strategies to evaluated the simulated effect on ICU survival (Supplementary Methods), according to subject-matter knowledge. The risk ratios (RRs) and risk differences (RDs) and their 95% CIs were used to compare the cumulative risk between the natural course and different fluid management strategies. The covariates for adjustment were selected because they were potential confounders based on subject-matter knowledge. Because of the rapidly declined number of patients over than two months, we set a truncated endpoint as 42 days after CCI diagnosed to avoid unstable estimation (Supplementary Methods, Figs. S35 and S36).

Ethics committee approval
All data from patients were retrospectively collected from the electronic healthcare records systems (in form

---

Medical Information Mart for Intensive Care III; eICU-CRD – eICU Collaborative Research Database; AUMC – Validation II cohort (AmsterdamUMCdb dataset); EHR – Electronic healthcare records; cart - Classification and regression trees; midas – Weighted predictive mean matching; pmm - Predictive mean matching; rf - Random Forest imputations; XGBoost – eXtreme Gradient Boosting.

of third-party public databases or hospital healthcare systems) which originated from daily clinical work. All data was performed in de-identification before the analysis. This study was approved by the local ethics committees (2021DZSKT-YBB-016). The third-party public databases used in this study (MIMIC-IV, MIMIC-III, eICU-CRD, and AmsterdamUMCdb) were exempted from our institutional review board approval. No informed consent was obtained, and all available data in the databases were anonymous.

Statistical analysis

Kaplan–Meier survival analysis was employed to estimate survival within ICU stay and evaluated by the log-rank test. In survival analysis, death events during ICU stay were set as the endpoints, and ICU length of stay was set as the length of survival. The 28-day cumulative hazard during ICU stay after CCI diagnosis was calculated. We evaluated the correlation between the subphenotypes and SOFA score to check whether the subphenotype is explained by illness severity.[21] Comorbidities of each subphenotype were inspected. The continuous variables were firstly examined for Gaussian distribution by the Shapiro–Wilk test and then expressed as the mean (standard deviation) or median (interquartile range, IQR) as appropriate. For comparisons, we employed the *Kruskal–Wallis* test for continuous data and the *Chi-square* test for categorical data. A P value < 0.05 for two sides is considered statistical significance.

Software and codes

For data extraction: Structured Query Language (SQL) – Google Cloud BigQuery; PostegreSQL 14; The SQL codes used in this study are based on https://github.com/MIT-LCP/mimic-code; https://github.com/alistairewj/mimic-iv-aline-study[22]; https://github.com/MIT-LCP/eicu-code; https://github.com/nus-mornin-lab/oxygenation_kc[23,24]; https://github.com/AmsterdamUMC/AmsterdamUMCdb[16]; For data analysis: R [version 4.1.2 and 4.2.0); Platform: x86_64-w64-mingw32/x64 [64-bit]; Dataset integration, exploration, summary (lubridate, data.table, tidyverse, DataExplorer, tableone); Imputation and pre-processing (mice; Amelia); Clustering (OPTICS – dbscan, Consensus Kmeans – ConsensusClusterPlus, LPA – mclust, tidyLPA, SOM – kohonen, Hierarchical clustering – factoextra); Visualization (lattice, ggridges, uwot, circlize, fmsb, ggalluvial); Classification model (xgboost, Ckmeans.1d.dp, multiROC); The SHAP can be found at: https://github.com/slundberg/shap; Shiny application – shiny, rsconnect); Window application construction (R Portable, golem, shiny, Node.js, Electron). The Eletron can be found at: https://github.com/listen2099/electron-quick-start/archive/master.zip); Survival analysis (survival, survminer); Parametric G-formula (gfoRmula).

For more details, please see Supplementary Methods.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

PL and JR had accessed and verified the underlying data and had final responsibility for the decision to submit for publication.

Results

Patients and study cohorts

A total of 8145 patients were diagnosed with CCI in this study, with 3761 patients in the Derivation (MIMIC) cohort, 2136 patients in the Treatment (MIMIC-IV) cohort, 2987 patients in the Validation I (eICU-CRD, US) cohort, 1263 patients in the Validation II (AmsterdamUMCdb, Euro) cohort, and 134 patients in the Validation III (Jinling, CN) cohort (Supplementary Fig. S1). In all cohorts, the male had a more significant proportion of CCI. Indeed, senior patients were predominant overall in CCI. The overall ICU mortality is 18.3% in the Derivation cohort, and the median ICU duration is 20.6 days (IQR: 16.8–27.1). In the overall CCI population, the most common admission diagnoses were circulatory system disease (N = 1134 (30.4%)), injury and poisoning (N = 670 (17.8%)), infectious and parasitic diseases (N = 645 (17.2%)), respiratory system disease (N = 521 (13.9%)) (Supplementary Table S3). In general, patients with CCI had a greater amount of poor disposition than good disposition (Derivation and Validation I, P < 0.01).

Derivation of subphenotypes showed heterogeneous characteristics

Consensus clustering identified 4-classes CCI subphenotypes named the Phenotype A, B, C, and D (Supplementary Fig. S8). Comparisons of characteristics and clinical outcomes were shown (Table 1, Fig. 2A–B, Supplementary Tables S3 and S4).

Compared to the overall CCI populations and the other phenotypes, patients of Phenotype A were the subgroup with the relatively mildest illness (with the lowest SOFA score and the youngest age), and they tended to have the highest survival. Patients of Phenotype B took the largest proportion of all patients (33.3%) with the oldest age, significantly higher bicarbonate, and lower anion gap. Indeed, we also found that they had the lowest WBC count (median [IQR]: 10.5 [7.9,13.0]), which was the lowest among the overall population and the other phenotypes. Patients of Phenotype C were more severe than A and B with multiple organ dysfunction encountering hypernatremia and hyperchloremia. Furthermore, they also tended to have a higher urea-creatinine ratio, representing a potentially

Characteristics	Overall	Phenotype				P value
		A	B	C	D	
Patient number	3761	938	1254	774	795	
Demographics						
Sex (%)						0.011
Male	2171 (57.7)	545 (58.1)	679 (54.1)	469 (60.6)	478 (60.1)	
Female	1590 (42.3)	393 (41.9)	575 (45.9)	305 (39.4)	317 (39.9)	
Age, years (median [IQR])	65 [53,75]	50 [38,61]	71 [62,79]	69 [60,78]	65 [54,74]	<0.0001
Ethnicity (%)						0.21
Asian	81 (2.2)	19 (2.0)	28 (2.2)	12 (1.6)	22 (2.8)	
Black	333 (8.9)	82 (8.7)	104 (8.3)	59 (7.6)	88 (11.1)	
Hispanic	122 (3.2)	31 (3.3)	34 (2.7)	28 (3.6)	29 (3.6)	
Other/Unknown	762 (20.3)	207 (22.1)	258 (20.6)	147 (19.0)	150 (18.9)	
White	2463 (65.5)	599 (63.9)	830 (66.2)	528 (68.2)	506 (63.6)	
Unit type of admission (%)						<0.0001
CCU	631 (16.8)	100 (10.7)	188 (15.0)	137 (17.7)	206 (25.9)	
CSRU	254 (6.8)	65 (6.9)	89 (7.1)	65 (8.4)	35 (4.4)	
MICU	1003 (26.7)	244 (26.0)	335 (26.7)	197 (25.5)	227 (28.6)	
NICU	182 (4.8)	63 (6.7)	76 (6.1)	37 (4.8)	6 (0.8)	
SICU	1030 (27.4)	265 (28.3)	375 (29.9)	182 (23.5)	208 (26.2)	
TSICU	661 (17.6)	201 (21.4)	191 (15.2)	156 (20.2)	113 (14.2)	
Admission type (%)						<0.0001
Elective[a]	560 (14.9)	122 (13.0)	194 (15.5)	118 (15.2)	126 (15.8)	
Emergency	3201 (85.1)	816 (87.0)	1060 (84.5)	656 (84.8)	669 (84.2)	
Comorbidities						
Elixhauser comorbidity index (mean (SD))	9.3 (7.8)	9.6 (8.0)	8.9 (7.5)	9.2 (7.9)	9.9 (8.1)	0.36
Charlson comorbidity index (mean (SD))	5.8 (2.9)	3.5 (2.5)	6.2 (2.5)	6.3 (2.6)	7.0 (2.9)	<0.0001
Cardiovascular (%)	2163 (57.5)	457 (48.8)	749 (59.7)	451 (58.3)	506 (63.6)	<0.0001
Neurologic (%)	934 (24.8)	276 (29.5)	327 (26.1)	198 (25.6)	133 (16.7)	<0.0001
Diabetes (%)	1052 (28.0)	202 (21.6)	339 (27.0)	227 (29.3)	284 (35.7)	<0.0001
Respiratory (%)	978 (26.0)	188 (20.1)	353 (28.1)	194 (25.1)	243 (30.6)	<0.0001
Renal (%)	706 (18.8)	87 (9.3)	186 (14.8)	129 (16.7)	304 (38.2)	<0.0001
Peptic ulcer (%)	99 (2.6)	13 (1.4)	29 (2.3)	21 (2.7)	36 (4.5)	0.0001
Hepatic (%)	623 (16.6)	118 (12.6)	145 (11.6)	133 (17.2)	227 (28.6)	<0.0001
Malignant (%)	352 (9.4)	57 (6.1)	123 (9.8)	84 (10.9)	88 (11.1)	0.0007
Rheumatoid (%)	97 (2.6)	16 (1.7)	33 (2.6)	22 (2.8)	26 (3.3)	0.20
AIDS (%)	36 (1.0)	10 (1.1)	9 (0.7)	8 (1.0)	9 (1.1)	0.76
Infection and sepsis cases						
Suspected infection on day of CCI diagnosed (%)	809 (21.5)	216 (23.0)	224 (17.9)	160 (20.7)	209 (26.3)	<0.0001
Sepsis on day of CCI diagnosed (%)	785 (20.9)	207 (22.1)	212 (16.9)	157 (20.3)	209 (26.3)	<0.0001
SOFA (mean (SD))	6.1 (3.9)	5.2 (3.4)	5.3 (3.5)	6.0 (3.3)	8.8 (4.1)	<0.0001
Outcomes						
ICU length of stay, days (median [IQR])	20.6 [16.8, 27.1]	21.00 [16.7, 28.0]	19.5 [16.2, 26.0]	21.4 [17.0, 28.0]	21.0 [16.9, 28.2]	<0.0001
Hospital length of stay, days (median [IQR])	27.0 [21.0, 37.0]	28.00 [21.0, 37.0]	25.0 [20.0, 34.0]	27.0 [21.0, 37.0]	29.0 [21.0, 41.0]	<0.0001
ICU mortality (%)	689 (18.3)	121 (12.9)	186 (14.8)	146 (18.9)	236 (29.7)	<0.0001
Hospital mortality (%)	877 (23.3)	146 (15.6)	254 (20.3)	198 (25.6)	279 (35.1)	<0.0001
Hospital discharge location (%)						<0.0001
Good disposition (%)	1559 (41.5)	490 (52.2)	550 (43.9)	302 (39.0)	217 (27.3)	<0.0001
Home	105 (2.8)	51 (5.4)	34 (2.7)	7 (0.9)	13 (1.6)	
Home health care services	211 (5.6)	78 (8.3)	66 (5.3)	30 (3.9)	37 (4.7)	
Rehabilitation	1243 (33.0)	361 (38.5)	450 (35.9)	265 (34.2)	167 (21.0)	
Poor disposition (%)	2202 (58.5)	448 (47.8)	704 (56.1)	472 (61.0)	578 (72.7)	<0.0001
Another hospital	91 (2.4)	28 (3.0)	26 (2.1)	15 (1.9)	22 (2.8)	
Death	882 (23.5)	148 (15.8)	254 (20.3)	199 (25.7)	281 (35.3)	
Hospice	61 (1.6)	11 (1.2)	24 (1.9)	14 (1.8)	12 (1.5)	

(Table 1 continues on next page)

Characteristics	Overall	Phenotype				P value
		A	B	C	D	
(Continued from previous page)						
Long-term acute care facility	832 (22.1)	190 (20.3)	258 (20.6)	181 (23.4)	203 (25.5)	
Other/Unknown	13 (0.3)	5 (0.5)	5 (0.4)	2 (0.3)	1 (0.1)	
Skilled nursing facility	323 (8.6)	66 (7.0)	137 (10.9)	61 (7.9)	59 (7.4)	

CCI – chronic critical illness; IQR – interquartile range; SD – standard deviation; CCU – Coronary Care Unit; CSRU –Cardiac Surgery Recovery Unit; MICU – Medical Intensive Care Unit; SICU – Surgical Intensive Care Unit; TSICU – Trauma Surgical Intensive Care Unit; SOFA –sequential organ failure assessment; AIDS – acquired immunodeficiency syndrome; [a]In admission type, 'Elective' includes planned visiting, surgical admission, and observation status admission.

Table 1: Characteristics of the CCI phenotypes in the Derivation cohort.

hypercatabolic state. Patients with Phenotype D had the most critical multiple organ dysfunction, such as renal dysfunction (highest creatinine and BUN and lowest urine output), anaemia (lowest haemoglobin), hemodynamic instability (lowest mean blood pressure), coagulatory abnormality (lowest platelet count), and acidosis (highest anion gap and lowest bicarbonate). Indeed, patients of Phenotype D had the lowest urea-creatinine ratio among all phenotypes. Moreover, patients of Phenotype D had the highest proportion of infectious disease diagnosis at admission (survival: 121 (21.6%); non-survival: 65 (27.5%), Supplementary Table S3).

Furthermore, we investigated comorbidities of each subphenotype in the Derivation cohort, as shown by radar diagrams. Patients of Phenotype A had the fewest comorbidities. Patients of Phenotype D had the most comorbidities, significantly accompanied by diabetes, renal, hepatic, and cardiovascular complications. All subtypes of patients with CCI had a significant number of cardiovascular complications, while Phenotype D had the greatest proportion (Fig. 2E).

Validation, reproducibility, and sensitive analysis of CCI subphenotypes

To assess the reproducibility and stability of developed CCI subphenotypes from the Derivation cohort, we performed an external validation on the Validation I cohort by using the same consensus kmeans settings and obtained four clusters with similar features (Supplementary Figs. S10–S12; Table S5). Pairwise comparison of variables and UMAP analyses in CCI Phenotype A to D showed similarity between consensus kmeans, LPA, and hSOM, indicating the good reproducibility of CCI subphenotypes (Fig. 2C–D; Supplementary Figs. S23–S27; Table S8). The sensitive analyses obtained similar subphenotypes properties (Supplementary Methods; Figs. S9, S13–S22). Overall, these analyses confirmed the reliability of the CCI subphenotypes we derived.

Clinical outcomes of CCI subphenotypes

Among all CCI subphenotypes in the Derivation cohort, Phenotype A had the lowest ICU (12.9%) and hospital (15.6%) mortality, while Phenotype D had the highest ICU (29.7%) and hospital (35.1%) mortality (Table 1). In

Phenotype A from the Derivation cohort, patients tended to have a more significant proportion of good prognosis (52.2%). Phenotype D was the most severe subtype with the greatest proportion of poor prognosis (72.7%). Phenotype B and C had a less poor prognosis than Phenotype D (B: 56.1%; C: 61.0%). However, they had a larger proportion of destinations to skilled nursing facilities than Phenotype D (B: 10.9%; C: 7.9%; D: 7.4%).

We employed Kaplan–Meier survival analysis and calculated the cumulative hazard of 28-day survival after CCI was diagnosed during ICU stay (Fig. 3). The Kaplan–Meier curves showed that Phenotype A was the lowest group among all subphenotypes (P < 0.01), while Phenotype B, C, and D had worse short-term mortality during ICU stay. Similar outcomes were observed in the other validation cohorts (Supplementary Figs. S28 and S29; Tables S5–S7).

Relationship between CCI subphenotypes and organ dysfunction severity

We further inspected the relationship between subphenotypes and severity of illness, such as SOFA score, to explore whether the previously identified CCI subphenotypes were simply a reflection of the severity of organ dysfunction (Supplementary Fig. S30; Table S9). The alluvial plot showed that there was no complete direct correspondence between subtypes and SOFA groups (Supplementary Fig. S31). Hence, the derived subphenotypes cannot be simply interpreted by the severity of organ dysfunction.

Construction of CCI subphenotypes classifier

The simplified XGBoost classifier showed good effectiveness. We plotted the multiclass ROC curves and calculated their AUC (A: 0.903; B: 0.791; C: 0.900; D: 0.948; Macro: 0.885; Micro: 0.889) (Supplementary Fig. S32B-C) for effectness evaluation and SHAP value plots for model interpretation (Supplementary Fig. S33). We then applied the compact model to the Validation II and III cohort to obtain subphenotype memberships (Supplementary Fig. S34), and their subphenotype characteristics were consistent with the Derivation cohort (Supplementary Fig. S11; Tables S6 and S7). For
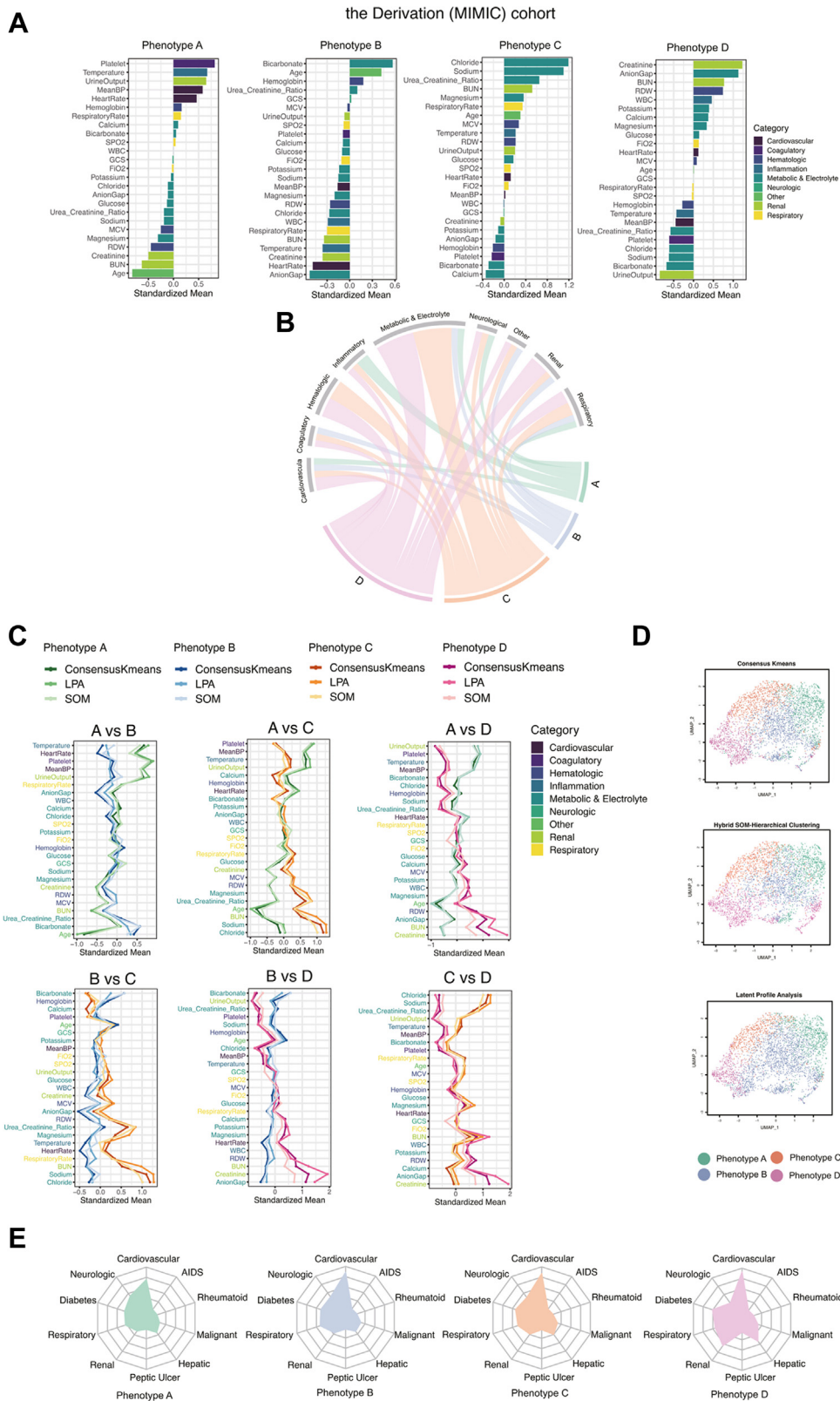
Fig. 2: Characteristics of CCI subphenotypes. The analysis of characteristics of four CCI subphenotypes in the Derivation (MIMIC) cohort. (A) Characteristic histograms of the four phenotypes ordered in the most remarkable clinical variables from the positive direction to the negative

the convenience of clinicians, we further built an interactive-interface application. By entering these six indicators, clinicians can easily obtain the appropriate subphenotype for a single patient to support decision-making (Supplementary Fig. S32D).

Heterogeneous treatment effect of daily fluid balance among subphenotypes

The parametric G-formula models for the overall population and each phenotype independently showed distinct patterns of the association between daily fluid balance and ICU mortality (Fig. 4, Supplementary Figs. S37 and S38; Tables S10–S15).

For the overall CCI population, we found that a bundle of daily fluid management strategy between −500 and 1000 mL was associated with decreased in truncated 42-days ICU mortality after CCI diagnosed, while the both excessive or even higher negative and positive balance ($\leq$−2500 mL or $\geq$ 2500 mL per day) seemed to be harmful for patients with CCI. Under different scaled daily fluid balance management strategies, patients of Phenotype A only were associated with increased ICU mortality when undertook aggressive daily positive balance (more than 3000 mL fluid net intake; crude risk (95% CI): 0.95 (0.56 to 1); risk ratio (95% CI): 2.64 (1.08 to 3.5)). For patients of Phenotype B, strategies of daily fluid balance from over −3000 to −1000 mL and 2000 to over 3000 mL were associated worse prognosis, while daily fluid balance between 0 and 500 mL showed lower ICU mortality. In patients of Phenotype C, daily fluid balance from over −3000 to −500 mL and 2500 to over 3000 mL were associated worse prognosis. A decrease in truncated 42-days (after CCI diagnosed) ICU mortality was observed in this subphenotype population only when the simulated daily positive fluid balance was between −500 and 1000 mL. As patients with Phenotype D, we observed that strategies of fluid management between over −3000 to −500 mL and over 3000 mL associated with increase ICU mortality. Only when daily positive fluid balance between 200 and 1500 mL in patients of Phenotype D showed improved in prognosis. The effects of other strategies in different CCI patient populations showed no statistically significances.

Discussion

In this study, we derived and validated four subphenotypes of patients with CCI based on five EHR datasets covering the US, Europe, and China populations using three unsupervised methods. Each subphenotype represented distinct patterns of clinical characteristics and outcomes. We found that these subphenotypes cannot be simply interpreted by the SOFA score. An easy-to-use classifier for subphenotype prediction was developed to facilitate clinical utility in time and accurately. Furthermore, we studied the heterogeneous treatment response of daily fluid balance among different CCI subphenotypes and found that for patients from different CCI phenotype, the beneficial threshold interval of daily fluid management was different.

CCI was defined as a patient with a long-term ICU stay with critical organ dysfunction.[6] Previous studies also reported such kinds of patients as persistent critical illness (PCI).[25] Notably, a specific type of patients with CCI with inflammatory and metabolic disorders was recognized as persistent inflammation, immunosuppression, and catabolism syndrome (PIICS).[10] Despite the dispute of naming, it is beyond doubt that prolonged ICU stay and critical organ dysfunction are the two most prominent features of patients with CCI. Management of patients with CCI is extremely challenging and empirical due to the significant heterogeneity of disease patterns, prognoses, and treatment responses.[9] Bagshaw et al. reported that the pathophysiological characteristics and behaviour of prolonged ICU hospitalized patients vary from acute critical cases.[26] Many attempts to apply treatment from current guidelines of primary diseases to patients with CCI have not achieved curative effect as expected or even failed,[9,27] which urges the academic community to further explore the precise and individualized recognition and treatment of CCI. Recently, the

---

direction. The bars were coloured in relative categories of each variable. (B) In chord diagrams, the ribbons connect from a specific phenotype to a system category if the group mean value was worse than the overall mean value for the entire Derivation (MIMIC) cohort. (C) The mean value of clinical variables for phenotyping from the Derivation (MIMIC) cohort were scaled from −1 to 1 as Z-score standardization (X-Axis). Each subset represented the comparison between each two subphenotypes. Variable's comparison in all panels ranked the differences from the greatest positive to the greatest negative direction. Each variable was assigned to the relative category of features by specific colour (Y-Axis). In all line plot panels, the standardized mean of each variable in each phenotype was shown in consensus kmeans, latent profile analysis, and hybrid SOM-hierarchical clustering represented by different colours. (D) Visualization of phenotypes using UMAP technique in the Derivation (MIMIC) cohort by consensus kmeans, latent profile analysis, and hybrid SOM-hierarchical clustering with coloured phenotypes distribution. (E) Radar plots show the relationship and comparison between CCI phenotypes and comorbidities in the Derivation (MIMIC) cohort. SOM – Self-organizing map; LPA – Latent profile analysis; BUN – blood urea nitrogen; FiO_2 – fraction of inspired oxygen; GCS – Glasgow coma scale; MCV – mean corpuscular volume; MeanBP – mean blood pressure; RDW – red cell distribution width; SPO_2 – oxygen saturation; WBC – white blood cell; AIDS – acquired immunodeficiency syndrome; MIMIC – Medical Information Mart for Intensive Care; UMAP – Uniform manifold approximation and projection.
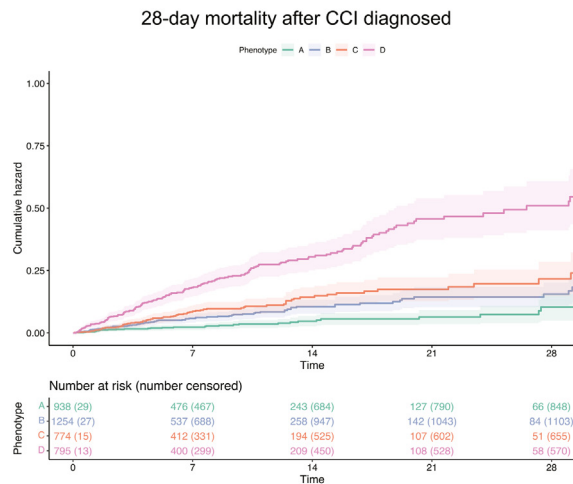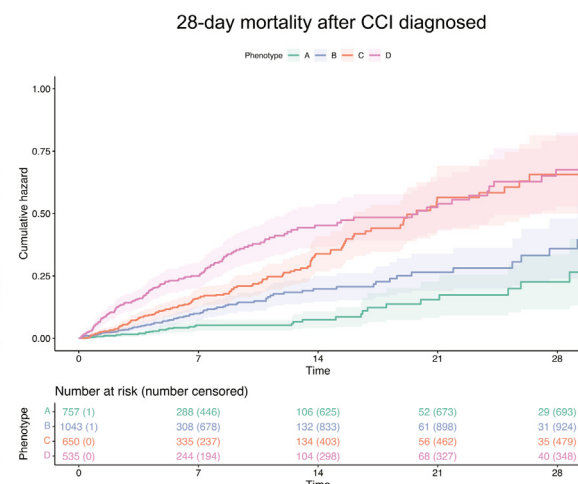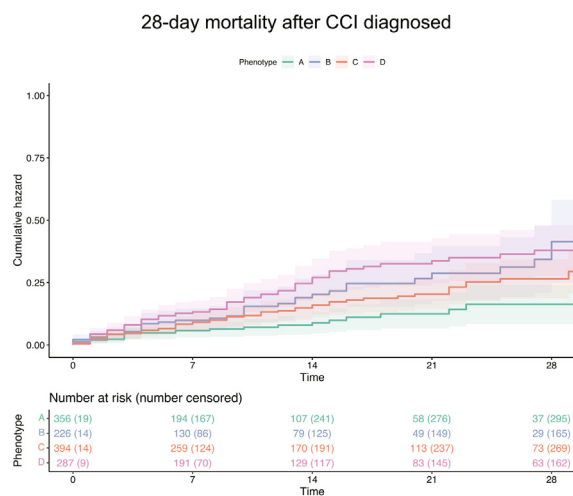
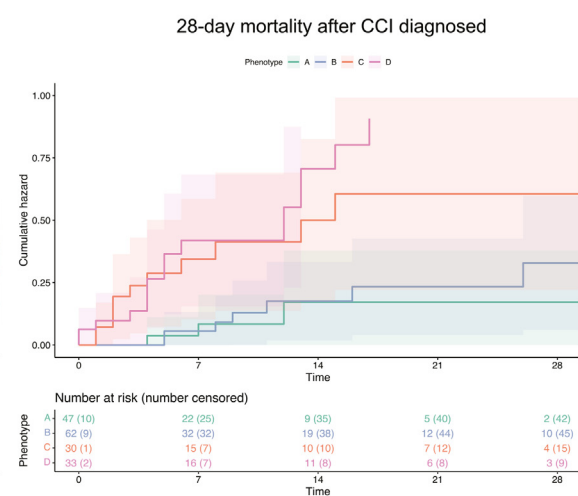**Fig. 3: Survival analysis for consensus Kmeans phenotyping in the Derivation cohort, Validation I cohort, Validation II cohort, and Validation III cohort during ICU stay.** In each subset figure, left side represented the Kaplan–Meier curve plots for four cohorts of 28-day mortality within ICU stay. The X-axis denotes the time (days) after patients were diagnosed with CCI (at Day 14 in ICU) and Y-axis denotes the cumulative hazard. CCI – Chronic critical illness; AUMC – AmsterdamUMCdb dataset.

development of unsupervised methods to discover disease phenotypes facilitated the implementation of precision medicine, which inspired our exploration for the heterogeneity of the CCI population.[28]

Notably, among all CCI subphenotypes, Phenotype A is the mildest illness subclass with the relatively youngest population, while Phenotype B was the subclass with the highest proportion of the elderly. Indeed, several studies reported that advanced age was

considered a risk factor for prognosis in CCI, which is consistent with our findings.[3] The majority of the young population might be explainable for the least organ dysfunction and lowest ICU mortality in patients of Phenotype A, though such group of patients was still critically ill compared to non-CCI patients. In comparison, the weakness of patients with advanced age might be a critical reason for the difference in prognosis between Phenotype A and B. In addition, patients of
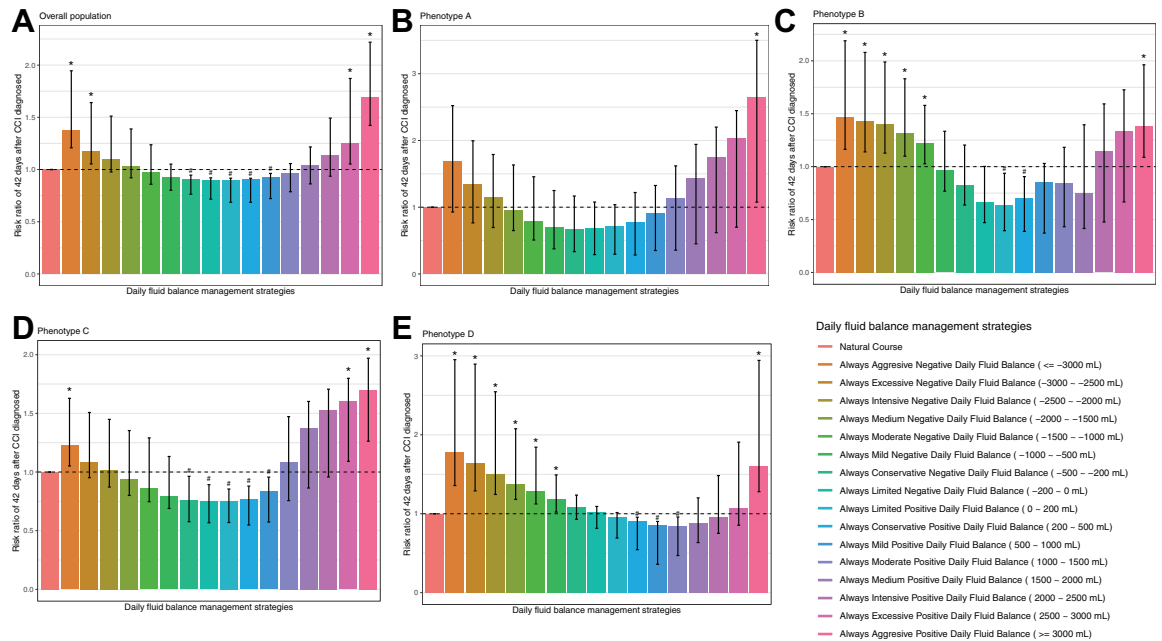
Fig. 4: Association between daily fluid management strategies and ICU mortality among the overall population and subphenotypes.
According to parametric G-formula estimation, A-E represented the risk ratio of different strategies compared to the natural course. A strategy that showed statistical significance with worsened prognosis was marked with an asterisk (*); while a strategy that showed statistical significance with improved prognosis was marked with a pound sign (#).

Phenotype B also showed alkaline status (low anion gap and high bicarbonate) and low WBC. Whereas patients of Phenotype D were the most severe type with high levels of anion gap, and low levels of bicarbonate, implying a potential acid status. Previous studies showed that a higher anion gap is correlated with a higher inflammatory marker (leukocyte count and CRP),[29,30] which were close to our findings of Phenotype B (low anion gap) and D (high anion gap). Moreover, patients of Phenotype C showed a specifically hyper-catabolic state, represented by the high urea-creatinine ratio.[18] Haines et al. developed this marker to assess the catabolic state in PCI after trauma, and Zhang et al. validated its effectiveness in PCI after sepsis.[18,31] Interestingly, patients of Phenotype D had a very low urea-creatinine ratio. Although patients with Phenotype C and D were the two subphenotypes with the most severe organ dysfunction and worst prognoses, they showed opposite metabolism-catabolism states. Therefore, it is necessary to distinguish these two subgroups of patients and give them appropriate interventions. According to these findings, there was a potential association between CCI subphenotypes we derived and PIICS which needs further exploration and validation. Notably, patients of Phenotype C also have hypernatremia and hyperchloremia. Rugg et al. studied the relationship between PIICS and ICU-acquired hypernatremia, which supports our finding and implies the correlated mechanism

between electrolytes and metabolism in CCI.[32] These findings were validated in the large sample size and geographically distinct populations and showed convincing robustness.

Interestingly, since CCI tends to be a continuation result of multiple acute critical illness, such as sepsis and acute respiratory depression syndrome (ARDS), the dynamic pathophysiological changes and the transition of disease states of patients during their ICU stay may be closely related and the occurrence and entry of CCI subphenotypes. Xu et al. reported a group of sepsis subtypes based on trajectory clustering, which revealed the alteration of their organ function levels over time.[33] Further studies exploring the evolutionary relationship between the CCI subphenotype we identified and the other acute critical illness with the related subphenotypes are warranted in the future.

Academic concern and interest in fluid therapy in critically ill patients have risen exponentially in recent years. Here, our study showed remarkable differences in water-electrolyte-related indicators in different CCI subphenotypes, suggesting a potential heterogeneous association between CCI subphenotypes on fluid therapy. Indeed, previous studies showed that conservative fluid therapy or de-resuscitation could lower in-hospital organ dysfunction or death.[34] The balance of body fluids can be influenced by many factors. For example, urine output, which is the main contributor to negative fluid

balance, is also affected by fluid infusion, blood transfusion, and other therapeutic measures. This forms a complex network of causal pathways, and conventional estimation methods are difficult to use to eliminate the interference of these confounding factors. Therefore, we included multiple covariates in a parametric G-formula model to estimate the different effect under distinct strategies of longitudinal daily fluid balance management. In the overall population and all CCI phenotypes, the relationship between daily fluid net amount and cumulative risk ratio compared to the natural course showed a U-shaped correlation, which was consistent with previous studies – either improper negative or positive fluid balance had adverse effect on pathophysiology in critically ill patients.[11] However, we found there were evident differences in the interval of net benefit for four CCI phenotypes under different treatment strategies, suggesting that the thresholds for management of daily fluid balance that needed to be received were different for different subtypes of patients with CCI. Patients of Phenotype A showed prognostic deterioration only under aggressive positive daily fluid balance strategy, considering the illness severity and age of such patients, we inferred that there may be two reasons of this finding. Firstly, patients of Phenotype A may have higher tolerance to fluid. Additionally, patients with relatively stable conditions seems less likely to received extreme fluid administration in clinical practice, thus our simulated estimation were not sufficient to derive significance of these strategies with the natural course, which may also imply that patients of Phenotype A did not need to pay special attention to the precise management of fluid balance. However, individualized fluid management for Phenotype B, C, and D patients is important since the beneficial intervals were different from the overall CCI populations and each other subphenotypes. Patients of Phenotype D seem to be particularly required to receive a positive fluid balance, which we inferred that may be because these patients were critically ill and unstable, thus they tended to require extra fluid supplementation to maintain the organ systems. Further investigations are necessary.

There were several limitations in this study. First, all data we used in the analysis were collected from the daily routine clinical work EHR system, which limited the selection of available variables for phenotyping and existing error records and missingness. Meanwhile, the validation dataset from China is relatively small. This may reduce the confidence in the results. We performed multiple procedures for reproducibility to ensure the reliability of the results. Moreover, this study included patients with CCI with a total relatively large sample size of 8145 patients, by using multiple validations and sensitivity analysis approaches, our findings performed well. Compared with other studies, our work avoided the limitations of a single centre or region,[6,17] and

subphenotypes were well validated in the US, Europe, and China cohorts. Second, the Treatment cohort was generated from retrospective data, and thus confounding factors may contribute to the unreliability of the results, and omission of unmeasured confounders may lead to a biased estimation. In parametric G-formula analysis, we did not include the information of chronic heart failure and daily sepsis status due to data availability, which requires follow-up researches by incorporating more updated research data to explore this issue. In addition, the different time period of ICU admission may reflect updates in clinical recognition and practices, especially for fluid management including infusion rate, type of fluid, measures of fluid negative balance (such as diuretic and CRRT strategies), and nutritional support, which could cause biased estimation. We analysed multiple covariates when building the parametric G-formula model, which increased the stability and credibility of the conclusions to the greatest extent. A carefully-designed randomized controlled trial is very necessary in the future. Third, many subphenotype identification systems were commonly challenging to present to clinicians, which limited their usage. We built an easy-to-use classifier and deployed it as an application, which directly facilitated clinical work promotion. Fourth, our study did not explore the latent mechanism of each subphenotype. It requires further exploration in future research incorporating multi-omics data and in-vitro/in-vivo experiments. Fifth, whether the subphenotypes derived based on the day of CCI diagnosis, are stable during a prolonged ICU stay needs to be further validated for characteristics and outcomes. The four subphenotypes of CCI that we identified were based on the indicators at the time of diagnosis of the corresponding patients, but whether there would be disturbances between subphenotypes in the subsequent course of the disease has not been explored, and this important question needs to be elaborated in future longitudinal clustering analysis, such as growth mixture modeling. Sixth, our study only focused on early inhospital prognoses, while long-term prognoses follow-up and life quality inspection are also crucial for CCI. A further study focuses on regarding long-term outcomes (6 months and longer) in CCI survivors' population across the distinct subphenotypes is needed. Seventh, in this study, we investigated the calculation of daily total fluid intake. However, we did not explore other aspects of fluid therapy such as the types of fluid therapy, infusion rates, strategies for negative fluid balance, and testing for patients' volume responsiveness. These areas require further investigation in future studies. Prospective clinical trials in the future are necessarily required.

In conclusion, we identified four data-driven phenotypes that demonstrated heterogeneous patterns among patients with CCI, regardless of geographically

different populations. Using the classifier applied, clinicians can easily recognize CCI subphenotypes at the bedside and in time. The four subphenotypes will deepen the understanding of pathophysiological features and guide individualized treatment of fluid balance in patients with CCI. Our findings highlight the promising prospect of precision management in patients with CCI and path the way to the improvement in clinical management and trial enrolment. Further prospective researches are required to confirm our findings in the utility of practice.

References
1   Nelson JE, Cox CE, Hope AA, Carson SS. Chronic critical illness. *Am J Respir Crit Care Med*. 2010;182(4):446–454.
2   Sakusic A, Gajic O. Chronic critical illness: unintended consequence of intensive care medicine. *Lancet Respir Med*. 2016;4(7):531–532.
3   Rosenthal MD, Kamel AY, Rosenthal CM, Brakenridge S, Croft CA, Moore FA. Chronic critical illness: application of what we know. *Nutr Clin Pract*. 2018;33(1):39–45.
4   Kahn JM, Le T, Angus DC, et al. The epidemiology of chronic critical illness in the United States*. *Crit Care Med*. 2015;43(2):282–287.
5   Ohbe H, Matsui H, Fushimi K, Yasunaga H. Epidemiology of chronic critical illness in Japan: a nationwide inpatient database study. *Crit Care Med*. 2021;49(1):70–78.
6   Gardner AK, Ghita GL, Wang Z, et al. The development of chronic critical illness determines physical function, quality of life, and long-term survival among early survivors of sepsis in surgical ICUs. *Crit Care Med*. 2019;47(4):566–573.
7   Stoller JK, Xu M, Mascha E, Rice R. Long-term outcomes for patients discharged from a long-term hospital-based weaning unit. *Chest*. 2003;124(5):1892–1899.
8   Li S, Wu X, Ren J. Diagnostic criteria for chronic critical illness should Be standardized. *Crit Care Med*. 2021;49(10).
9   Viglianti EM, Kruser JM, Iwashyna T. The heterogeneity of prolonged ICU hospitalisations. *Thorax*. 2019;74(11):1015–1017.
10 Mira JC, Gentile LF, Mathias BJ, et al. Sepsis pathophysiology, chronic critical illness, and persistent inflammation-immunosuppression and catabolism syndrome. *Crit Care Med*. 2017;45(2):253–262.
11 Balakumar V, Murugan R, Sileanu FE, Palevsky P, Clermont G, Kellum JA. Both positive and negative fluid balance may be associated with reduced long-term survival in the critically ill. *Crit Care Med*. 2017;45(8):e749–e757.
12 41st International Symposium on Intensive Care and Emergency Medicine. *Crit Care*. 2022;26(1):72.
13 Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV. PhysioNet; 2021. . Accessed March 16, 2021.
14 Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3(1):160035.
15 Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data*. 2018;5:180178.
16 Thoral PJ, Peppink JM, Driessen RH, et al. Sharing ICU patient data responsibly under the society of critical care medicine/European society of intensive care medicine joint data science collaboration: the Amsterdam university medical centers database (AmsterdamUMCdb) example*. *Crit Care Med*. 2021;49(6).
17 Darden DB, Brakenridge SC, Efron PA, et al. Biomarker evidence of the persistent inflammation, immunosuppression and catabolism syndrome (PICS) in chronic critical illness (CCI) after surgical sepsis. *Ann Surg*. 2021;274(4):664–673.
18 Haines RW, Zolfaghari P, Wan Y, Pearse RM, Puthucheary Z, Prowle JR. Elevated urea-to-creatinine ratio provides a biochemical signature of muscle catabolism and persistent critical illness after major trauma. *Intensive Care Med*. 2019;45(12):1718–1731.
19 Stortz JA, Mira JC, Raymond SL, et al. Benchmarking clinical outcomes and the immunocatabolic phenotype of chronic critical illness after sepsis in surgical intensive care unit patients. *J Trauma Acute Care Surg*. 2018;84(2):342–349.
20 McGrath S, Lin V, Zhang Z, et al. gfoRmula: an R package for estimating the effects of sustained treatment strategies via the parametric g-formula. *Patterns (N Y)*. 2020;1(3).
21 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.
22 Hsu DJ, Feng M, Kothari R, Zhou H, Chen KP, Celi LA. The association between indwelling arterial catheters and mortality in hemodynamically stable patients with respiratory failure: a propensity score analysis. *Chest*. 2015;148(6):1470–1476.
23 Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med*. 2018;24(11):1716–1720.
24 van den Boom W, Hoy M, Sankaran J, et al. The search for optimal oxygen saturation targets in critically ill patients: observational data from large ICU databases. *Chest*. 2020;157(3):566–573.
25 Iwashyna TJ, Hodgson CL, Pilcher D, et al. Timing of onset and burden of persistent critical illness in Australia and New Zealand: a retrospective, population-based, observational study. *Lancet Respir Med*. 2016;4(7):566–573.
26 Bagshaw SM, Stelfox HT, Iwashyna TJ, Bellomo R, Zuege D, Wang X. Timing of onset of persistent critical illness: a multi-centre retrospective cohort study. *Intensive Care Med*. 2018;44(12):2134–2144.
27 Rosenthal MD, Bala T, Wang Z, Loftus T, Moore F. Chronic critical illness patients fail to respond to current evidence-based intensive care nutrition secondarily to persistent inflammation, immunosuppression, and catabolic syndrome. *J Parenter Enter Nutr*. 2020;44(7):1237–1249.

28   Reddy K, Sinha P, O'Kane CM, Gordon AC, Calfee CS, McAuley DF. Subphenotypes in critical care: translation into clinical practice. *Lancet Respir Med*. 2020;8(6):631–643.

29   Farwell WR, Taylor EN. Serum anion gap, bicarbonate and biomarkers of inflammation in healthy individuals in a national survey. *CMAJ (Can Med Assoc J)*. 2010;182(2):137–141.

30   Zampieri FG, Kellum JA, Park M, et al. Relationship between acid-base status and inflammation in the critically ill. *Crit Care*. 2014;18(4):R154.

31   Zhang Z, Ho KM, Gu H, Hong Y, Yu Y. Defining persistent critical illness based on growth trajectories in patients with sepsis. *Crit Care*. 2020;24(1):57.

32   Rugg C, Strohle M, Treml B, Bachler M, Schmid S, Kreutziger J. ICU-acquired hypernatremia is associated with persistent inflammation, immunosuppression and catabolism syndrome. *J Clin Med*. 2020;9(9).

33   Xu Z, Mao C, Su C, et al. Sepsis subphenotyping based on organ dysfunction trajectory. *Crit Care*. 2022;26(1).

34   Silversides JA, Major E, Ferguson AJ, et al. Conservative fluid management or deresuscitation for patients with sepsis or acute respiratory distress syndrome following the resuscitation phase of critical illness: a systematic review and meta-analysis. *Intensive Care Med*. 2017;43(2):155–170.