7-20-2021

Phenotyping Cardiogenic Shock

Elric Zweck
Katherine L. Thayer
Ole K.L. Helgestad
Manreet Kanwar
Mohyee Ayouty

See next page for additional authors

Follow this and additional works at: https://scholarlycommons.henryford.com/cardiology_articles

Recommended Citation
Zweck E, Thayer KL, Helgestad OKL, Kanwar M, Ayouty M, Garan AR, Hernandez-Montfort J, Mahr C, Wenczer D, Sinha SS, Vorovich E, Abraham J, O’Neill W, Li S, Hickey GW, Josiassen J, Hassager C, Jensen LO, Holmvang L, Schmidt H, Ravn HB, Møller JE, Burkhoff D, and Kapur NK. Phenotyping Cardiogenic Shock. J Am Heart Assoc 2021; 10(14):e020085.

This Article is brought to you for free and open access by the Cardiology/Cardiovascular Research at Henry Ford Health Scholarly Commons. It has been accepted for inclusion in Cardiology Articles by an authorized administrator of Henry Ford Health Scholarly Commons.
Authors
Elric Zweck, Katherine L. Thayer, Ole K.L. Helgestad, Manreet Kanwar, Mohyee Ayouty, A. Reshad Garan, Jaime Hernandez-Montfort, Claudius Mahr, Detlef Wencker, Shashank S. Sinha, Esther Vorovich, Jacob Abraham, William O’Neill, Song Li, Gavin W. Hickey, Jakob Josiassen, Christian Hassager, Lisette O. Jensen, Lene Holmvang, Henrik Schmidt, Hanne B. Ravn, Jacob E. Møller, Daniel Burkhoff, and Navin K. Kapur
Phenotyping Cardiogenic Shock

Elric Zweck, MD; Katherine L. Thayer, MPH; Ole K. L. Helgestad, MD, PhD; Manreet Kanwar, MD; Mohyee Ayouty, MSc; A. Reshad Garan, MD; Jaime Hernandez-Montfort, MD; Claudius Mahr, MD; Detlef Wencker, MD; Shashank S. Sinha, MD; Esther Vorovich, MD; Jacob Abrahm, MD; William O’Neill, MD; Song Li, MD; Gavin W. Hickey, MD; Jakob Josiassen, MD; Lisette O. Jensen, MD, PhD, DMSci; Lene Holmvang, MD, DMSci; Henrik Schmidt, MD, DMSci; Hanne B. Ravn, MD, PhD, DMSci; Jacob E. Møller, MD, PhD, DMSci; Daniel Burkhoff, MD, PhD; Navin K. Kapur, MD

BACKGROUND: Cardiogenic shock (CS) is a heterogeneous syndrome with varied presentations and outcomes. We used a machine learning approach to test the hypothesis that patients with CS have distinct phenotypes at presentation, which are associated with unique clinical profiles and in-hospital mortality.

METHODS AND RESULTS: We analyzed data from 1959 patients with CS from 2 international cohorts: CSWG (Cardiogenic Shock Working Group Registry) (myocardial infarction [CSWG-MI; n=410] and acute-on-chronic heart failure [CSWG-HF; n=480]) and the DRR (Danish Retroshock MI Registry) (n=1069). Clusters of patients with CS were identified in CSWG-MI using the consensus k means algorithm and subsequently validated in CSWG-HF and DRR. Patients in each phenotype were further categorized by their Society of Cardiovascular Angiography and Interventions staging. The machine learning algorithms revealed 3 distinct clusters in CS: "non-congested (I)"; "cardiorenal (II)," and "cardiometabolic (III)" shock. Among the 3 cohorts (CSWG-MI versus DRR versus CSWG-HF), in-hospital mortality was 21% versus 28% versus 10%, 45% versus 40% versus 32%, and 55% versus 56% versus 52% for clusters I, II, and III, respectively. The "cardiometabolic shock" cluster had the highest risk of developing stage D or E shock as well as in-hospital mortality among the phenotypes, regardless of cause. Despite baseline differences, each cluster showed reproducible demographic, metabolic, and hemodynamic profiles across the 3 cohorts.

CONCLUSIONS: Using machine learning, we identified and validated 3 distinct CS phenotypes, with specific and reproducible associations with mortality. These phenotypes may allow for targeted patient enrollment in clinical trials and foster development of tailored treatment strategies in subsets of patients with CS.

Key Words: cardiogenic shock ■ clusters ■ heart failure ■ hemodynamics ■ machine learning ■ myocardial infarction ■ phenotypes
As a result, attempts at staging CS have been based on expert opinions and consensus.\textsuperscript{11–14} To avoid complexity, some of these classification systems include only a few variables and rely on specific, although arbitrary, cutoffs that introduce bias and fail to capture the full variability of patient profiles. The recently proposed Society of Cardiovascular Angiography and Interventions (SCAI) staging provides discriminatory potential for morbidity and mortality.\textsuperscript{15,16} It can be used to track the severity of shock over the course of a hospital stay.

Thus, a means of appropriately phenotyping patients with CS at admission remains a much-needed critical step in the development of treatment algorithms and prospective clinical studies to improve patient outcomes. New insights to subclassify patients with CS may be gained by using an unbiased, algorithmic approach to data analysis using machine learning (ML). Clustering algorithms (a form of unsupervised ML) have been effectively used for classification and phenotyping of other clinical syndromes and diseases, including diabetes mellitus, HF with preserved ejection fraction, and sepsis.\textsuperscript{17–20}

The objective of this investigation was to identify and evaluate CS phenotypes in large, contemporary CS data sets. We hypothesized that the use of ML algorithms can mathematically reduce routine clinical information at the time of presentation to discrete, reproducible phenotypes of CS. These phenotypes could further our understanding of shock physiology, inform patient selection for clinical trials, and be incorporated into clinical practice as an enhancement to our approach to risk assessment.

**METHODS**

**Overview**

This study involved 2 data sets and a multistep statistical approach. The CSWG (Cardiogenic Shock Working Group Registry) includes data from patients with both AMI-CS (CSWG-MI) and acute-on-chronic systolic HF (CSWG-HF), whereas the DRR (Danish Retroshock MI Registry) gathers data on patients with AMI-CS. The CSWG-MI cohort data were used as a derivation cohort for the CS phenotypes using consensus k-means clustering applied to relevant clinical variables. We then assessed phenotype reproducibility in the DRR data set with further subsequent validation in both the DRR and the CSWG-HF cohort. We assessed the association of phenotypes with mortality and clinical profiles across both causes: AMI and HF. Last, we studied the association between ML-derived phenotypes and in-hospital mortality within individual SCAI stages.

**Data Sources and Study Populations**

The CSWG is a multicenter database initiated in 2016 and currently includes 16 clinical sites across the United States, contributing data on patients with CS.\textsuperscript{15} For the purpose of this analysis, all available data by the 8 initial participating sites at the time of initial presentation for patients with CS between years 2016 and
2019 were included. Patients admitted to respective hospitals’ catheterization laboratories and intensive care units were screened by the clinical coordinators (retrospectively) if they met the predefined criteria for CS. The screening processes were physician adjudicated, and any patient who met the inclusion criteria was included. The registry includes a standardized set of data elements that were predefined by principal investigators and collected retrospectively. These include patient demographics, clinical presentation, procedural factors, and hospital characteristics. Patients from the CSWG with myocardial infarction (MI) as an underlying cause of shock represented our derivation cohort (CSWG-MI), whereas those with acute-on-chronic HF serve as a validation cohort (CSWG-HF). Quality assurance was achieved through monitoring at each site by the respective clinical coordinators and principal investigator. Values were centrally audited and screened by the CSWG research team for any discrepancies or major outliers and resolved with the submitting site. Data collection at each US clinical site and data sharing were approved by individual site’s institutional review boards. The need for informed consent was waived because of the retrospective deidentified nature of data collection. The CSWG is housed and analyzed by the lead operational team (E.Z., M.A., and K.T.) at Tufts Medical Center.

The DRR database is derived from the DNPR (Danish National Patient Registry), which records all patient contacts within the Danish healthcare system. For this study, baseline data on patients with AMI-CS hospitalized between 2012 and 2017 were collected retrospectively from 2 tertiary academic facilities in Denmark (Odense University Hospital and Copenhagen University Hospital Rigshospitalet) that cover nearly two thirds of the entire Danish population.1 In all cases, data were collected from medical records after being evaluated for the diagnosis of CS. The DRR was approved by the Danish Patient Safety Authority (file number: 3-3013-1133/1) and the Danish Data Protection Agency (file numbers: 16/7381 and 18/23756). The need for informed consent was waived. Further information on the data sources can be found in previous publications.1,15

**Definition of CS**

CS diagnosis was physician adjudicated at each site based on criteria defined by the CSWG.15 CS was defined retrospectively if at least 1 of the following 3 conditions were met: (1) a sustained episode of systolic blood pressure ≤90 mm Hg for at least 30 minutes or the need for vasoactive agents to maintain such blood pressure; (2) a cardiac index <2.2 L/min per m² determined to be secondary to cardiac dysfunction, in the absence of hypovolemia; or (3) the use of a temporary mechanical circulatory support device for suspected CS. Only adult (aged ≥18 years) patients with a known clinical outcome and sufficiently complete data for common variables in both data sets were considered for analyses. All data were recorded at an available time point as close to index hospital admission as possible.

**Data Processing and Variable Handling**

To derive the phenotypes, we first assessed the candidate variables’ distribution, missingness, and correlation (Data S1). Within the CSWG cohorts, variables and patients with high proportions of missing data were removed from the derivation data set, to ensure that overall missingness did not exceed 10% for imputation (Data S1). Any remaining missing values were imputed with random forest imputation. For the validation DRR cohort, only patients with complete data for clustering were included, to exclude any bias introduced by imputation. To rule out any potential bias introduced because of the exclusion step, we performed sensitivity analyses of CSWG-HF and DRR without removal of any patients. To reduce variable collinearity and to limit complexity and dimensionality, we used a classification algorithm to select variables from continuous clinical and laboratory data (Data S1).21 Concretely, we used a random forest classifier to identify demographic and laboratory variables that predicted in-hospital mortality in the CSWG-MI derivation cohort because it does not assume linear relationships between variables and removed highly correlated variables (|r|>0.6), leaving 6 variables to be included in the final analyses (Data S1).

**Cluster Analyses and Validation**

Clustering is an ML technique used to identify homogeneous subgroups within data, such that data points in each cluster are as similar as possible while being as different from other clusters as possible. K-means clustering is one of the most common unsupervised algorithms that iteratively tries to partition the data set into distinct, nonoverlapping subgroups (clusters) where each data point belongs to only one group. We deployed semisupervised consensus k-means clustering on the 6 identified variables with highest predictive value using the ConsensusClusterPlus package22 for R 3.6.0 and defined the optimal number of clusters (k) using several metrics (Data S1). We selected various methods to visualize the clusters (t-distributed stochastic neighbor embedding plots,23 chord diagrams, and rank plots) and their characteristics which are detailed in the Supplemental Material.

For validation, first, cases in the DRR were clustered independently using the same variables and cluster algorithm as in the derivation cohort to assess external reproducibility. Phenotype characteristics were then scrutinized within and between these cohorts to assess whether a similar pattern can be seen when the 2
cohorts are clustered completely independently. Finally, cases in DRR and CSWG-HF were assigned to a respective cluster using the centroids from the clusters in the CSWG MI derivation cohort (Data S1) to validate applicability of the clusters in external data sets and future clinical practice. Distribution, mortality, and characteristics of the clusters in the validation cohorts were compared with those in the derivation CSWG-MI cohort.

Application of SCAI Shock Classification Scheme
To put the novel phenotypes into clinical context, we assessed the SCAI shock stages in the CSWG cohort and combined this established method of CS classification to our proposed phenotypes. Each patient was categorized into the most severe SCAI stage encountered during his/her hospital stay, using a method previously described by Thayer et al. Of note, the phenotypes were derived from data at the time of initial presentation, whereas the SCAI staging was denoted as the "highest stage" during hospitalization.

Statistical Analysis
P values were calculated using ANOVA, Kruskal-Wallis test, or χ² test. Continuous data are displayed as mean and SD or as median and interquartile ranges, depending on distribution. All statistical tests were performed using SAS Enterprise Guide 9.4 (SAS Institute Inc, Cary, NC) and Python 3.7.4 (Python Software Foundation, DE). Figures were created using Python 3.7.4 and GraphPad Prism 8.2.1 (GraphPad Software, Inc, San Diego, CA). Statistical significance threshold was P<0.05.

RESULTS
A total of 1959 patients (410 in CSWG-MI cohort, 1069 in DRR cohort, and 480 in CSWG-HF cohort) were eligible for final analyses (Figure S1). Mean ages at presentation were 65±13, 66±11, and 57±14 years, whereas male sex represented 68%, 77%, and 75% patients in the CSWG-MI, DRR, and CSWG-HF cohorts, respectively (Table 1). The in-hospital mortality from CS was noted to be 39% in CSWG-MI cohort, 45% in DRR cohort, and 26% in CSWG-HF cohort. The AMI-CS cohorts (CSWG-MI and DRR) were similar on various clinical variables but differed in age and treatment strategy, especially in use of temporary mechanical circulatory support (Table 1). CSWG-HF differed from the MI cohorts on several clinical features, including younger age, higher body weight, more severe kidney and liver injury, and lower lactate levels (Table 1).

The 8 clinical variables that were most associated with in-hospital mortality were: glomerular filtration rate, serum bicarbonate, serum lactate, alanine aminotransferase, platelet count, serum creatinine, white blood cell count, and blood urea nitrogen levels (Figure S2). Of these, glomerular filtration rate, lactate, serum bicarbonate, and alanine aminotransferase were the most predictive ones (Figure S2).

Phenotypes in CS
Consensus k-means clustering in the CSWG-MI derivation cohort identified 3 as the optimal number (k) of clusters, based on the calculated silhouette score, cluster consensus, and other metrics (Figure 1 and Figure S3). These clusters demonstrated important differences beyond the variables used to design them, suggesting that the clustering algorithm successfully identified 3 discrete clusters of CS that were able to identify as distinct clinical phenotypes. On the basis of their clinical characteristics, we labeled the 3 phenotypes as "noncongested (I)," "cardiorenal shock (II)," and "cardiometabolic shock (III)."

The 3 phenotypes differed from one another on various demographic and clinical parameters (Table 2). Radar plots were used to display the deviation of metabolic and hemodynamic values from the mean value (derived from the total derivation cohort average values) (Figure 2), and chord plots were used to reveal the sources of phenotype differences (Figure S4). The noncongested phenotype (I) exhibited lower heart rate, filling pressures (right atrial and pulmonary capillary wedge pressures), and a higher blood pressure relative to the other phenotypes. This represents a relatively stable profile of a noncongested patient with CS. In contrast, the patients in the cardiorenal shock (II) group were older, with multiple comorbidities. They exhibited a lower heart rate, elevated pulmonary arterial and pulmonary capillary wedge pressures, as well as lower glomerular filtration rate, suggesting renal involvement from shock. Last, the patients in the cardiometabolic shock (III) group exhibited elevated lactate, alanine aminotransferase, heart rate, and right atrial pressure, along with low blood pressure, cardiac power output, and index. This suggested a multiorgan involvement, featured by transaminases and lactic acidosis in a patient with CS.

Validation of Phenotypes
First-pass validation of the derived phenotypes was performed to test external reproducibility by using the same consensus k-means clustering algorithm de novo in the DRR validation cohort. Using this approach, similar patterns of the clusters were identified compared with the derivation cohort (Figures S5 and S6). In the second validation procedure, we assigned patients from DRR and CSWG-HF to the derived phenotypes using a classifier based on the
### Table 1. Cohort Characteristics

| Characteristics | CSWG-MI Cohort | DRR Cohort | CSWG-HF Cohort |
|-----------------|---------------|-----------|---------------|
|                 | No. (%)       | No. (%)   | No. (%)       |
| Nonsurvivors    | 161 (39.27)   | 478 (44.71)| 127 (26.46)  |
| Men             | 279 (68.05)   | 819 (76.61)| 362 (75.42)  |
| IABP            | 249 (60.73)   | 127 (11.88)| 200 (41.67)  |
| ECMO            | 129 (31.22)   | 43 (4.02)  | 93 (19.38)   |
| Impella         | 170 (41.46)   | 152 (14.22)| 109 (22.71)  |
| Any t-MCS       | 395 (96.34)   | 288 (26.99)| 325 (67.71)  |
| No t-MCS        | 15 (3.66)     | 799 (73.01)| 155 (32.29)  |
| Multiple t-MCS  | 134 (32.68)   | 34 (3.19)  | 76 (15.83)   |
| Mechanical ventilation | 242 (59.02) | Not captured | 163 (33.96) |
| Vasopressor/inotrope use | 318 (77.56) | Not captured | 384 (80)     |
| Vasodilators    | 62 (15.12)    | Not captured | 201 (41.88) |
| History of hypertension | 281 (68.54) | 512 (47.9)   | 216 (45)     |
| History of CKD (any stage) | 73 (17.8) | Not captured | 169 (35.21) |
| History of COPD | 24 (5.85)     | 104 (9.73) | 49 (10.21)   |
| History of CVA/TIA | 57 (13.9)  | 85 (7.95)  | 76 (15.83)   |
| Prior HF        | 92 (22.44)    | Not captured | 363 (75.63) |
| Prior MI        | 110 (26.83)   | 152 (14.22)| 132 (27.5)   |
| History of PCI  | 138 (33.66)   | Not captured | 94 (19.58)   |
| History of CABG | 33 (8.05)     | Not captured | 45 (9.38)    |
| History of diabetes mellitus | 179 (43.66) | 173 (16.18)| 142 (29.58) |
| History of PVD  | 20 (4.88)     | 77 (7.2)   | 18 (3.75)    |
| MeanSD          | Mean | SD | MeanSD | Mean | SD | MeanSD | Mean | SD |
| Age, y          | 65.12 | 13.27 | 66 | 11.04 | 56.87 | 14.38 |
| Weight, kg      | 81.9 | 18.81 | 81.11 | 15.77 | 86.69 | 22.24 |
| Laboratory Values (Normal Range) | Mean | SD | % OOR | Mean | SD | % OOR | Mean | SD | % OOR |
| Sodium, mEq/L   | 137.16 | 4.2 | 137.88 | 4.52 | 134.22 | 5.53 |
| Potassium, mEq/L (3.6–5.1) | 4.3 | 0.73 | 23.82 | 4.07 | 38.64 | 4.26 |
| HCO\(_3\), mEq/L (21–28) | 20.2 | 4.79 | 55.00 | 18.47 | 73.71 | 4.60 |
| BUN, mg/dL (6–24) | 27.83 | 16.65 | 55.40 | 24.6 | 33.84 | 37.86 |
| Creatine, mg/dL (0.6–1.3) | 1.61 | 0.18 | 50.64 | 1.48 | 44.43 | 1.84 |
| WBC, 10\(^3\)/mm\(^3\) (4–11) | 14.59 | 7.09 | 64.19 | 16.39 | 81.01 | 10.77 |
| Hemoglobin, g/dL (11–16) | 12.52 | 2.58 | 39.34 | 13.54 | 21.68 | 11.98 |
| Hematocrit, % (32–47) | 37.33 | 7.53 | 34.81 | 39.04 | 31.59 | 36.79 |
| Platelets, 10\(^9\)/mm\(^3\) (150–400) | 217.94 | 88.94 | 24.79 | 247.27 | 89.01 | 198.95 |
| ALT, U/L (<55) | 173.17 | 392.18 | 43.49 | 198.81 | 464.87 | 237.83 |
| Total bilirubin, mg/dL (0.2–1.1) | 0.88 | 0.55 | 23.97 | 0.77 | 19.15 | 1.78 |
| INR (0.9–1.3) | 1.38 | 0.47 | 67.83 | 1.27 | 21.12 | 1.95 |
| GFR, mL/min per 1.73 m\(^2\) (>90) | 55.37 | 26.83 | 88.49 | 58.29 | 89.71 | 50.48 |
| Lactate, mEq/L (0.5–2.2) | 4.54 | 3.99 | 35.46 | 5.76 | 81.67 | 3.94 |
| pH (7.35–7.45) | 7.28 | 0.15 | 68.20 | 7.26 | 82.76 | 7.35 |
| MAP, mm Hg | 74.72 | 16.71 | 64.09 | 11.79 | 73.06 | 13.14 |

(Continued)
Zweck et al Phenotyping Cardiogenic Shock

cluster centroids of the derivation cohort to test for potential future clinical application (Figure S7). Both validation cohorts revealed a similar trend in mortality and comparable clinical profiles across the 3 clusters, as was seen in the independent clustering of the CSWG-MI and DRR cohorts (Table S1 and Figure S8). These trends were consistent in all sensitivity analyses, including the imputation of all missing values in CSWG-HF and DRR instead of removing cases with a large number of missing values (Tables S2 and S3 and Figure S9).

### Association of Phenotypes With Outcomes

In-hospital mortality rates differed among phenotypes but were similar within a given phenotype, even across cohorts. Relative to phenotype I, patients in phenotype II had a higher mortality (CSWG-MI odds ratio [OR], 3 [95% CI, 1.8–5.1]; DRR OR, 2 [95% CI, 1.4–2.8]), whereas those in phenotype III (CSWG-MI OR, 4.6 [95% CI, 2.7–5.1]; DRR OR, 3.5 [95% CI, 2.5–4.9]) were at highest risk of dying (Figure 3).

### Phenotypes and Shock Stages

We compared the distribution of each phenotype’s patients with regard to their most severe SCAI shock stage encountered during the hospital stay (Figure 4). The risk of developing stage D or E shock during the hospital stay was lowest in phenotype I and highest in phenotype III for patients with both CS-MI and CS-HF. Within each phenotype, the SCAI staging (C–E) further stratified mortality. Similarly, within each SCAI stage, the 3 phenotypes further stratified mortality. In a bivariate logistic regression model, both SCAI stage (C–E) and phenotype were significant predictors (P<0.0001) of in-hospital mortality in CSWG-MI, CSWG-HF, and CSWG-MI and CSWG-HF combined (Table S4).

### DISCUSSION

This is the first study using ML approaches in large, multicenter cohorts of patients with CS. Using this novel approach, we identified 3 distinct clusters of CS patient profiles in our derivation cohort of patients with MI-CS from the CSWG. Next, we tested the reproducibility of our analysis in an independent, Danish registry of patients with MI-CS in patients with HF-CS from the CSWG and identified the same 3 clusters with matching rates of in-hospital mortality. These CS phenotypes exhibited distinct demographic, hemodynamic, and metabolic signatures and correlated with inpatient mortality. The “noncongested” (phenotype I) group represented lowest mortality, and
the "cardiometabolic shock" (phenotype III) group resulted in highest mortality across all cohorts. We further identified that these 3 clusters were directly associated with mortality within individual SCAI stages. Accordingly, these findings may improve risk stratification and enable the development of treatment algorithms tailored to each phenotype of CS and inform patient selection in future clinical trials in CS.

Major discriminators of phenotype membership included glomerular filtration rate, lactate, serum bicarbonate, and alanine aminotransferase, which have all previously been associated with risk of mortality in CS.\textsuperscript{24,25} Previously, SCAI staging was shown to have a strong correlation with mortality in a heterogeneous intensive care unit population.\textsuperscript{16} Patients with refractory shock (SCAI shock stage E) were shown to have a significantly higher in-hospital mortality than other stages, regardless of cause.\textsuperscript{16} Identifying the clinical CS phenotypes that impact outcomes could allow for early classification into treatment groups. In our analysis, SCAI stage C had the lowest and stage E, the highest, mortality in each individual phenotype cluster. These findings support that the 3 phenotypes of CS reflect clinically relevant features that are expected to be associated with mortality. Therefore, the phenotypes seem to be compatible with the SCAI staging system and provide further support to it in guiding even more individualized therapy. Unlike the SCAI staging system, whose aim is to characterize disease severity as it evolves over the course of a hospital stay, the aim of the current analysis was the identification of CS phenotypes at the time of admission.

![Figure 1. Derivation of the clusters: consensus k-means clustering.](image)

**Figure 1.** Derivation of the clusters: consensus k-means clustering.

A, Representative plots illustrating the method of consensus k-means clustering in the CSWG (Cadiogenic Shock Working Group Registry) myocardial infarction derivation cohort. Each column represents one patient, whereas each row displays the assigned clusters. Well-defined (ie, segregated) squares indicate stable clusters. Compared with k (number of clusters)=2 and k=4, k=3 shows highest cluster stability, suggesting that 3 may be a good choice for the number of clusters. B, A t-distributed stochastic neighbor embedding (TSNE) plot for visual representation of the clusters in a 2-dimensional space. The algorithm uses probability estimates to calculate the similarity of data points in the high-dimensional space (ie, identifies the “neighbors”) and then calculates the distance of these “neighbors” in a lower-dimensional space (in this case, 2 dimensions).\textsuperscript{23} The wider the different clusters separate in the plot, the larger is the difference between them.

| Characteristics | Cluster/Phenotype I “Noncongested” CS | Cluster/Phenotype II “Cardiorenal” CS | Cluster/Phenotype III “Cardiometabolic” CS |
|-----------------|--------------------------------------|--------------------------------------|------------------------------------------|
| Mean age, y     | ≈60                                  | ≈70                                  | ≈65                                     |
| Comorbidities   | Few                                  | DM2, CKD, hypertension...            | Few                                     |
| Blood pressure  | ↓                                    | ↓                                    | ↓                                       |
| Congestion      | None                                 | Left ventricular                     | Right ventricular                        |
| Heart rate      | ↔                                    | ↔                                    | ↑†                                     |
| Hemoglobin      | ↔                                    | ↓                                    | ↔                                       |
| Transaminases   | ↔                                    | ↔                                    | ↑†                                     |
| Lactate         | ↔ or ↑                               | ↓                                    | ↑†                                     |
| Kidney function | ↔                                    | ↓                                    | ↓                                       |

CKD indicates chronic kidney disease; CS, cardiogenic shock; and DM2, type 2 diabetes mellitus.
Although “labeling” phenotypes with 1-dimensional titles (noncongested, cardiorenal, and cardiometabolic) may obscure the fascinating mathematical interactions of the entirety of variables defining the phenotypes, each phenotype was composed of strikingly different characteristics: young patients with few comorbidities, normal hematocrit, low lactate, and good kidney function in phenotype I; older patients with chronic kidney disease, diabetes mellitus, hypertension, and anemia in phenotype II; and patients with lactic acidosis, inflammation, and elevated liver enzymes in phenotype III. Hemodynamically, phenotype I was characterized by relatively high cardiac output and blood pressures; phenotype II exhibited pulmonary congestion (high wedge pressure); phenotype III displayed systemic congestion (high central venous pressure but low wedge pressure, suggesting involvement of right-sided HF), high heart rate, and low cardiac output and arterial blood pressure. Taken together, these findings suggest that phenotype I represents the noncongested patient with high likelihood of salvage. Phenotype II, on the other hand, may involve more left-sided HF with worsening kidney function (cardiorenal), whereas phenotype III represents worsening venous congestion, liver damage, and likelihood of multiorgan involvement. Identifying these distinct phenotypes may change the way we classify patients with CS, each of whom may benefit from a different management strategy that could be tested in prospective trials.

Reproducibility and external validity are important measurements of cluster quality. A strength of this study is the finding that similar phenotypes evolved when we clustered 2 completely independent AMI-CS cohorts on the same variables, especially as these cohorts represent patients from 2 countries on different continents with different healthcare systems and treatment strategies. Future studies may validate these phenotypes globally, even further supporting the universal existence of 3 distinct groups of CS with sufficiently nonoverlapping characteristics and prognosis, such that patients can be classified on presentation. In addition, many of the phenotypes’ clinical characteristics align with existing clinical consensus, including the concepts that...
venous congestion and high lactate are associated with increased mortality, making their incorporation into practice more feasible.

The phenotypes of CS may also allow for the stratification of patients in clinical trials, both past and future, which could improve their design and interpretation. The inability of randomized controlled trials in CS to demonstrate significant differences in outcomes, despite technological advances and improvement in hemodynamic markers, implicates heterogeneity as a potential confounder. Previous work with ML in sepsis (another clinically heterogeneous population) has shown clinical phenotypes within large, randomized trials derive harm or benefit from an intervention, which may be opposite from the larger population when all patients are grouped together.18

Although the 3 phenotypes of CS were derived from patients with AMI-CS, these were equally applicable to patients with HF-CS. Despite baseline differences between the HF-CS and AMI-CS cohorts, the clinical presentation and mortality trends of the phenotypes were consistent, regardless of whether they were derived from patients with HF-CS or AMI-CS. This finding indicates applicability in both the 2 most common forms of CS.

Creating phenotypes in CS through ML has the potential to address both implicit biases and identify subgroups within large populations that may derive harm or benefit from a certain therapeutic intervention. Information gained from this type of data analysis may also impact future clinical trials. Although we found similar treatment protocols across all phenotypes in our analysis, the vast differences in mortality we report between these phenotypes stress the importance of more granularity in defining patient characteristics and phenotypes in future treatment protocols and clinical trial designs.

**Limitations**

Major limitations of this study include the retrospective nature of the data sources, limited number of patients and clinical variables, and the inability to assess long-term outcomes. ML algorithms tend to perform better on larger data sets, and the sample size may limit the complexity of a cluster model. However, there is currently a lack of large, comprehensive, contemporary multicenter CS databases that represent all causes of CS. Within these limitations, we were able to demonstrate and validate reproducible, clinically relevant phenotypes in CS using ML algorithms, which were associated with mortality. Future studies may collect comprehensive data in a prospective manner and may allow for enhanced, even more nuanced examination of the CS phenotypes.

In addition, the clusters in this article are semisupervised. Operating on the assumption that variables driving mortality are clinically the most relevant variables in CS, we identified important variables using supervised ML based on mortality association before applying an unsupervised clustering algorithm (semisupervised learning). However, because the phenotypes differed in several clinical features, even beyond the variables included in the clustering algorithm, this is likely to be the case for characteristics not captured in our data sets as well, thus providing a comprehensive classification. The selected ML algorithm, (consensus) k means, is a well-established algorithm in unsupervised learning,17,18 particularly...
in medical research, but is not simply applicable to categorical data; thus, only continuous variables were used for clustering. Renal impairment had a major impact on cluster assignment but did not discriminate between acute and chronic kidney injury. Furthermore, although it may be a limitation that the calendar time frames of data collection for the CSWG and DRR cohorts were not identical, they do overlap and changes in CS treatment protocol were not significant enough to introduce any bias in our results. Although we were able to assess the maximum SCAI stage across the patients’ hospital course, based on drug and device escalation, the initial SCAI stage at admission could not be validly identified retrospectively because there was only admission data for most patients in CSWG, and for patients to be assigned SCAI stage D or E, they would have to fail to respond to initial interventions first.\textsuperscript{15} Taken together, these limitations identify the need for new, larger, prospective CS registries that account for temporal changes in patient variables and treatment to improve our understanding of the nature of these phenotypes.

in medical research, but is not simply applicable to categorical data; thus, only continuous variables were used for clustering. Renal impairment had a major impact on cluster assignment but did not discriminate between acute and chronic kidney injury. Furthermore, although it may be a limitation that the calendar time frames of data collection for the CSWG and DRR cohorts were not identical, they do overlap and changes in CS treatment protocol were not significant enough to introduce any bias in our results. Although we were able to assess the maximum SCAI stage across the patients’ hospital course, based on drug and device escalation, the initial SCAI stage at admission could not be validly identified retrospectively because there was only admission data for most patients in CSWG, and for patients to be assigned SCAI stage D or E, they would have to fail to respond to initial interventions first.\textsuperscript{15} Taken together, these limitations identify the need for new, larger, prospective CS registries that account for temporal changes in patient variables and treatment to improve our understanding of the nature of these phenotypes.

CONCLUSIONS
We report that using an ML approach to define unbiased clusters in the complex and heterogeneous clinical syndrome of CS is feasible. Our analysis of multicenter cohorts of patients with CS identified 3 distinct phenotypes of CS with unique clinical profiles that are associated with different risks of in-hospital mortality. These clusters exhibited clinically relevant differences in hemodynamic and metabolic profiles and are equally applicable to patients with CS attributable to MI or acute-on-chronic HF. Future studies are needed to evaluate the clinical implications of these phenotypes as they relate to prognosis and optimal treatment.

ARTICLE INFORMATION
Received December 15, 2020; accepted May 17, 2021.

Affiliations
The CardioVascular Center, Tufts Medical Center, Boston, MA (E.Z., K.L.T., N.K.K.); Medical Faculty, Heinrich Heine University, Düsseldorf, Germany (E.Z.); Department of Cardiology, Odense University Hospital, Odense, Denmark (O.K.H., L.O.J., J.E.M.); Odense Patient Data Explorative Network, University of Southern Denmark, Odense, Denmark (O.K.H., J.E.M.);
Department of Cardiovascular Medicine, Allegheny Health Network, Pittsburgh, PA (M.K.); Tufts University School of Medicine, Boston, MA (M.A.); Beth Israel Deaconess Medical Center, Boston, MA (A.R.G.); Cleveland Clinic Florida, Weston, FL (J.H.); University of Washington Medical Center, Seattle, WA (C.M., S.L.); Baylor Scott & White Advanced Heart Failure Clinic, Dallas, TX (D.W.); Inova Heart and Vascular Institute, Falls Church, VA (S.S.S.); Northwestern Medicine, Chicago, IL (E.V.); Providence Heart Institute, Portland, OR (J.A.); Henry Ford Hospital, Detroit, MI (W.O.); UPMC Heart and Vascular Institute, Pittsburgh, PA (G.W.H.); Department of Cardiology, Rigshospitalet, Copenhagen, Denmark (J.J., C.H., L.H.); Department of Clinical Medicine, University of Copenhagen, Denmark (C.H., L.H., H.B.R.); Department of Cardiothoracic Anesthesia, Odense University Hospital, Odense, Denmark (H.S.); Department of Cardiac Anesthesiology, Rigshospitalet, Copenhagen, Denmark, (H.B.R.); and Cardiovascular Research Foundation, New York, NY (D.B.).

Sources of Funding
This work was supported by a NIH RO1 grant to NKK (RO1HL139785-01) and institutional grants from Abbott Laboratories Inc (Abbott Park, IL), Abiomed Inc (Danvers, MA), Boston Scientific Inc (Minneapolis, MN), and Getinge Inc (Wayne, NJ) to Tufts Medical Center and by a grant from the Danish Heart Foundation (16-R1107-A6576) to OKLH.

Disclosures
Dr Garan is an unpaid consultant for Abiomed Inc. Dr Hernandez-Montfort is a consultant for Abiomed Inc (research and education). Dr Burkhoff reports an unresolved, educational grant from Abiomed Inc to Cardiovascular Research Foundation. Dr Vorovich is a consultant and in the speakers’ bureau of Abiomed Inc. Dr Abraham is a consultant for Abbott Laboratories Inc (Wayne, NJ) to Tufts Medical Center and by a grant from the Danish Heart Foundation (16-R1107-A6576) to OKLH.

Supplementary Material
Data S1
Tables S1–S4
Figures S1–S9
References S26–30

REFERENCES
1. Helgestad OKL, Jøsassen J, Hasager C, Jensen LO, Holmvang L, Sorensen A, Frydland M, Lassen AT, Udesen NLJ, Schmidt H, et al. Temporal trends in incidence and patient characteristics in cardiogenic shock following acute myocardial infarction from 2010 to 2017: a Danish cohort study. Eur J Heart Fail. 2019;21:1370–1378. DOI: 10.1002/ejhf.1568.
2. Kolte D, Khera S, Aronow WS, Mujib M, Palaniswamy C, Sule S, Jain D, Gotswis D, Ahmed A, Frishman WH, et al. Trends in incidence, management, and outcomes of cardiogenic shock complicating ST-elevation myocardial infarction in the United States. J Am Heart Assoc. 2014;3:e000590. DOI: 10.1161/JAHA.113.000590.
3. van Diepen S, Katz JN, Albert NM, Henry TD, Jacobs AK, Kapur NK, Khera S, Aronow WS, Mujib M, Palaniswamy C, Sule S, Jain D, Gotswis D, Ahmed A, Frishman WH, et al. Trends in incidence, management, and outcomes of cardiogenic shock complicating ST-elevation myocardial infarction in the United States. J Am Heart Assoc. 2014;3:e000590. DOI: 10.1161/JAHA.113.000590.
4. Thiele H, Ohman EM, de Waha-Thiele S, Zeymer U, Desch S, Bolli R, S embry C, Niedermaier S, Deuschl T, et al. Contemporary management of cardiogenic shock: a scientific statement from the American Heart Association. Circulation. 2017;136:e232–e268. DOI: 10.1161/CIR.0000000000000525.
5. Thiele H, Ohman EM, de Waha-Thiele S, Zeymer U, Desch S, Management of cardiogenic shock complicating myocardial infarction: an update 2019. Eur Heart J. 2019;40:2671–2683. DOI: 10.1093/eurheartj/ehz563.
6. Thiele H, Jobs A, Ouweneel DM, Henriques JPS, Seyfarth M, Desch S, Eitel I, Poss J, Fuernau G, de Waha-Thiele S. Percutaneous short-term active mechanical support devices in cardiogenic shock: a systematic review and collaborative meta-analysis of randomized trials. Eur Heart J. 2017;38:3523–3531. DOI: 10.1093/eurheartj/ehx363.
7. Werdan K, Gießen S, Ebel H, Hochmair JS, Meier JJ, Col J, et al. Early revascularization in acute myocardial infarction complicated by cardiogenic shock: Shock Investigators: should we emergently revascularize occluded coronary arteries for cardiogenic shock. N Engl J Med. 1999;341:625–634. DOI: 10.1056/NEJM199908263410901.
8. Becker PM, Schräge B, Sinning CR, Schneck R, Schumacher S, Wäldecker C, Lindner D, Seifert M, Neumann JT, et al. Venosternal extracorporeal membrane oxygenation for cardiopulmonary support. Circulation. 2018;138:2298–2300. DOI: 10.1161/CIRCULATIONAHA.118.036691.
9. Burkhoff D, Garan AR, Kapur NK. The SCAI cardiogenic shock staging system gets taken for a test drive. J Am Coll Cardiol. 2019;74:2129–2131. DOI: 10.1016/j.jacc.2019.08.1020.
10. Baran DA, Grines CL, Bailey S, Burkhoff D, Hall SA, Henry TD, Hollenberg SM, Kapur NK, O’Neill W, Ormato JP, et al. SCAI clinical expert consensus statement on the classification of cardiogenic shock: this document was endorsed by the American College of Cardiology (ACC), the American Heart Association (AHA), the Society of Critical Care Medicine (SCCM), and the Society of Thoracic Surgeons (STS) in April 2019. Catheter Cardiovasc Interv. 2019;94:29–37. DOI: 10.1002/cct.28329.
11. Killip T, Kimball JT. Treatment of myocardial infarction in a coronary care unit. Am J Cardiol. 1967;20:457–464. DOI: 10.1016/0002-9149(67)90023-9.
12. Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary artery disease. N Engl J Med. 1979;300:1350–1358. DOI: 10.1056/NEJM197906140300502.
13. Stevenson LW, Pagani FD, Young JB, Jessup M, Miller L, Kormos RL, Nefte D, Ulissery K, Desvignes-Nickens P, Kirklin JK. INTERMACS profiles of advanced heart failure: the current picture. J Heart Lung Transplant. 2009;28:535–541. DOI: 10.1016/j.healun.2009.02.015.
14. Thayer KL, Zweck E, Ayouty M, Garan AR, Hernandez-Montfort J, Mayr C, Morine KJ, Newman S, Jorde L, Haywood JL, et al. Invasive hemodynamic assessment and classification of in-hospital mortality risk among patients with cardiogenic shock. Circ Heart Fail. 2020;13:e007099. DOI: 10.1161/CIRCHEARTFAILURE.120.007099.
15. Jentzer JC, van Diepen S, Barsness GW, Henry TD, Menon V, Rihal CS, Nauid SS, Baran DA. Cardiogenic shock classification to predict mortality in the cardiac intensive care unit. J Am Coll Cardiol. 2019;74:2117–2128.
16. Ahlqvist E, Storm P, Käräjämäki A, Martinell M, Dorkhan M, Carlsson A, Vikman P, Prasad RB, Aly DM, Almgren P, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol. 2018;6:381–386. DOI: 10.1016/S2213-8587(17)30052-1.
17. Seymour CW, Kennedy JN, Wang S, Chang C-C, Elliott CF, Xu Z, Berry S, Clermont G, Cooper G, Gomez H, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. JAMA. 2019;321:2003–2017. DOI: 10.1001/jama.2019.5791.
18. Zaharia OP, Strassburger K, Strom A, Bönhof GJ, Karusheva Y, Antoniou S, Bódis K, Markgraf DF, Burkart V, Müssig K, et al. Risk of diabetes-associated diseases in subgroups of patients with recent-onset diabetes: a 5-year follow-up study. Lancet Diabetes Endocrinol. 2019;7:684–694. DOI: 10.1016/S2213-8587(19)30187-1.
19. Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorgheide M, Bonow RO, Huang QC, Deo RC. Phenotyping for novel classification of heart failure with preserved ejection fraction. Circulation. 2015;131:269–279. DOI: 10.1161/CIRCUITAH.114.010637.
20. Altman RB, Ashley EA. Using "big data" to dissect clinical heterogeneity. Circulation. 2015;131:232–233. DOI: 10.1161/CIRCUITAHNA.A114.014106.
22. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26:1572–1573. DOI: 10.1093/bioinformatics/btq170.

23. van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008;9:2579–2605.

24. Pöss J, Köster J, Fuernau G, Etel I, de Waha S, Ouarrak T, Lassus J, Harjola V-P, Zeymer U, Thiele H, et al. Risk stratification for patients in cardiogenic shock after acute myocardial infarction. J Am Coll Cardiol. 2017;69:1913–1920. DOI: 10.1016/j.jacc.2017.02.027.

25. Harjola V-P, Lassus J, Sionis A, Keber L, Tarvastmäki T, Spinar J, Parissis J, Banaszewski M, Silva-Cardoso J, Carubelli V, et al. Clinical picture and risk prediction of short-term mortality in cardiogenic shock. Eur J Heart Fail. 2015;17:501–509. DOI: 10.1002/ejhf.260.

26. Stekhoven DJ, Buhrmann P. Missforest–non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28:112–118. DOI: 10.1093/bioinformatics/btr597.

27. Shah AD, Bartlett JW, Carpenter J, Nicholas O, Hemingway H. Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. Am J Epidemiol. 2014;179:764–774. DOI: 10.1093/aje/kwt1312.

28. Trunk GV. A problem of dimensionality: a simple example. IEEE Trans Pattern Anal Mach Intell. 1979;PAMI-1:306–307. DOI: 10.1109/TPAMI.1979.4768926.

29. Dolnicar S. A review of unquestioned standards in using cluster analysis for data-driven market segmentation, CD conference proceedings of the Australian and New Zealand marketing academy conference 2002 (ANZMAC 2002), Deakin University, Melbourne, 2-4 December 2002. Available at: https://ro.uow.edu.au/commpapers/273.

30. Formann AK. Die latent-class-analyse: einführung in theorie und anwendung. Weinheim [W. Germany]: Beltz; 1984. Available at: http://catalog.hathitrust.org/api/volumes/oclc/16851238.html
Supplements

Data S1. Supplemental Methods

Imputation and Data Processing

Imputation was carried out in the CSWG-MI and CSWG-HF cohorts using the MissForest function from the missingpy package for Python. Random forest imputation has been shown to perform well in epidemiological datasets and is able to deal with categorical variables. Before random forest (RF) imputation, we removed variables and patients with high missingness (cutoffs: more than 40% missingness for CSWG-MI and 34% for CSWG-HF, which was picked to ensure that overall missingness in the dataset did not exceed 10%) from the derivation dataset to ensure that our results are not driven by patients or variables with most missingness. Hemodynamic variables were gathered for patients in the CSWG but excluded from cluster analysis because the full set of hemodynamics was solely available in the CSWG registry (with a degree of missingness that was acceptable for descriptive statistics but insufficient for clustering which requires complete data), but not in the DRR registry thus using them for clustering would have prevented us from thoroughly validating the clusters. Hemodynamics are also hardly imputable in patients with all hemodynamic variables missing.

We performed a sensitivity analyses of the clustering results by deriving the clusters from CSWG MI datasets imputed with five different random seeds. After imputation, outliers, defined by adapted Tukey’s criteria (>3 interquartile ranges away from the 1st or 3rd quartile), were removed from further analyses and visual representation (Supplemental Figure 1). Variables with log-normal distributions were log-transformed. For cluster analyses, all continuous variables were normalized to the minimum of 0 and the maximum of 1. Only for the parallel coordinate plots and radar plots, variables were subsequently standardized to a mean of 0 and an SD of 1.

Variable selection

Correlating (non-orthogonal) variables can distort clustering, as several algorithms tend to weigh these variables higher than orthogonal variables, so they are important to identify and remove when running these analyses. Furthermore, especially in small datasets, clustering on too many variables can add too much granularity to the algorithm without achieving model generalizability. While no strict threshold exists to identify the optimal
number of variables to cluster on for a specific number of cases, a possible hint may be deduced from latent class analysis, where similar dimensionality problems occur\textsuperscript{29}: Formann proposed the minimal sample size to include no less than \(2^n\) cases (\(n = \) number of variables), preferably \(5*2^n\).\textsuperscript{30} For datasets of approximately 400-500 cases, this suggests that the maximum number of variables to cluster on is eight, preferably six variables. Therefore, to eliminate non-orthogonal variables and appropriately reduce the dimensionality of our model, we employed a classification algorithm to variable selection.

Based on the assumption, that variables driving mortality are clinically the most interesting variables in CS, we identified important variables by supervised ML based on mortality association prior to applying an unsupervised clustering algorithm (semi-supervised learning). We used a random forest classifier to predict in-hospital mortality in 10 bootstrapped samples of 75\% of the CSWG-MI derivation cohort. Unlike most regression models the random forest classifier does not assume linear relationships between variables. We identified mortality-predicting variables as variables with the highest average predictive importance in the 10 samples using the RandomForestClassifier function from sklearn.ensemble for Python. In a first run we used all continuous variables (including clinical and laboratory data) that remained after preprocessing independent of their correlation (Supplemental Figure 2). We then trained the random forest classifier again after removal of correlating variables and identified the six most predictive ones for the actual clustering process. Of note, based on the lab values collected, abnormal renal function did not discriminate between acute and chronic kidney injury.

**Clustering Procedures**

Before clustering, variables were normalized to a minimum of 0 and a maximum of 1. Distance measurement for all clustering algorithms was Euclidean distance. Consensus \(k\) means clustering was performed on 1000 bootstrap samples of the whole cohorts of the size of 80\% of the overall cohort in CSWG-MI and DRR. Consensus clustering provides several benefits including comprehensive cluster visualization, assessment of cluster stability, and estimation of an optimal \(k\) (number of clusters).\textsuperscript{18, 22} The optimal \(k\) was determined as the \(k\) achieving highest consensus of the derivation cohort samples, as well as using the Silhouette score, the Calinski-Harabasz criterion, the Davies-Bouldin index and the elbow method for \(k\) Means clustering in the total derivation cohort. For sensitivity analyses we tested if \(k\), cluster
consensus, and cluster distribution remained stable when (a) between five and eight variables were chosen for clustering instead of the identified six and (b) when imputed datasets with five different random seeds were used.

To be of clinical use, clusters would need assignable to patients individually without de-novo clustering of a full cohort.\textsuperscript{17, 20} We validated the applicability of our cluster assignment in individual patients using the centroids of the clusters in the derivation cohort to assign clusters to patients in the DRR and CSWG-HF validation cohorts to their respective nearest centroid using the NearestCentroid classifier from the sklearn.neighbors package. Composition of phenotypes and outcomes was compared in the different cohorts to externally validate the phenotypes gathered by the classifier.
Figure S1: Flow Chart of Study Populations and Data Processing

CSWG: Cardiogenic Shock Working Group (Registry); DRR: Danish Retroshock Registry; MI: Myocardial Infarction; HF: Heart Failure
Figure S2: Variable Importance in Random Forest Classifier

A Random Forest Classifier was trained on in-hospital mortality in the derivation cohort to identify the most mortality-driving variables. A and C: Variable importance was calculated as average importance of a variable in the random classifier in 10 runs with different seeds. Importance of the most predictive variable was set to 100%, and the others relative to this variable. A) shows the result using all variables (including correlating variables). Out of the most predictive variables, the correlating (i.e. “non-orthogonal”) variables were identified using a correlation matrix (B). In pairs of correlating (|r|>0.6) variables the variable with lower predictive value than the respective other variable was removed. The result is shown in C: The six variables with the highest predictive importance were the same in both instances, before and after removal of the non-orthogonal variables. ALT: Alanine Aminotransferase; BUN: Blood urea nitrogen; Crea: Serum Creatinine; GFR (CKDEPI): Glomerular Filtration Rate; Hgb: Hemoglobin; INR: International Normalized Ratio; WBC: White Blood Cell Count.
Figure S3: Specifying the Optimal Number of Clusters (k) from the CSWG-MI Derivation Cohort

A: Cluster-Consensus Plot showing the cluster-consensus values of clusters at each k. High values indicate cluster stability. B: Cumulative Distribution Function (CDF) plot for each k to determine where the CDF reaches a maximum without expense of consensus. Higher and “flatter” curves are favorable. C: Tracking plot for each k showing the cluster assignment of each case. Changing colors within a column indicate unstable cluster assignment, as these samples are changing clusters often in repeated runs. D: Different metrics for the quality of the clustering to determine the optimal k. Unlike A-C, this panel depicts the scores/criteria for k Means clustering on the full cohort and not the consensus k Means clusters. The scores determine, how well the variables entered in the clustering are clustered with different k. Higher silhouette scores and lower Davies-Bouldin scores indicate better clustering and relatively higher Calinski-Harabasz scores estimate the optimal k. Finally, the “elbow” in the elbow plot can be used to estimate the optimal k. The Silhouette, the Davies-Bouldin criterion and the elbow plot indicate an optimal k of 3 clusters. The Calinski-Harabasz suggests an optimal k at 2 or 3.
| Table S1: Patient Characteristics in the Clusters of CSWG-MI, DRR and CSWG-HF |
|---------------------------------------------------------------|
| Table displays only non-imputed data. Mean (SD) or n (%). ALT: Alamine Aminotransferase; BUN: Blood urea nitrogen; CI: Cardiac Index; CO: Cardiac Output; COPD: Chronic obstructive pulmonary disease; CPI: Cardiac Power Index; CPO: Cardiac Power Output; CSWG-HF: Cardiogenic Shock Working Group registry Heart Failure cohort; Creatinine: Serum creatinine; CSWG-MI: Cardiogenic Shock Working Group registry Myocardial Infarction cohort; CVA/TIA: Cerebrovascular accident/Transient Ischemic Attack; DBP: Diastolic blood pressure; DM2: Type 2 Diabetes Mellitus; DRR: Danish Retroschok Registry; ECMO: Extracorporeal membrane oxygenation; GFR: Glomerular Filtration Rate; Hgb: Hemoglobin; HTN: Hypertension; IABP: Intra-Aortic Balloon Pump; INR: International Normalized Ratio; LVEDD: Left ventricular end-diastolic dimension; MAP: Mean arterial pressure; PA Sat: Pulmonary Artery Saturation; PADP: Pulmonary Artery Diastolic Pressure; PAP: Pulmonary Artery Pressure; PAPI: Pulmonary Artery Pulsatility Index; PAP: Pulmonary Artery Systolic Pressure; PCWP: Pulmonary Capillary Wedge Pressure; PVD: Peripheral vascular disease; RAP: Right atrial pressure; SBP: Systolic blood pressure; WBC: White Blood Cell Count. |

|                      | CSWG MI | DRR INDEPENDENTLY CLUSTERED | DRR WITH ASSIGNED CLUSTERS | CSWG HF |
|----------------------|---------|-----------------------------|---------------------------|---------|
| N                    | N (%)   | N (%)                       | N (%)                     | N (%)   |
| Non-Survivors        | 30      | 20.55                       | 101 29.36                 | 90      |
| Male                 | 103     | 70.55                       | 271 78.78                 | 247 77.43 |
| IABP                 | 89      | 60.96                       | 44 12.79                  | 43 13.49  |
| ECMO                 | 36      | 24.66                       | 8 2.33                    | 10 3.13   |
| Impella              | 55      | 37.67                       | 40 11.63                  | 37 11.6   |
| Mechanical ventilation | 77     | 52.74                       | 79 19.65                 | 77 19.64  |
| Vasoconstr./Vasoexp  | 98      | 67.12                       | 329 95.64                 | 390 94.67 |
| Vasoconstr.          | 23      | 15.75                       | 129 37.5                  | 132 41.38 |
| History of HTN       | 84      | 57.53                       | 30 8.72                   | 27 8.46   |
| History of CKD (any stage) | 2   | 1.37                        | 24 6.98                   | 23 7.21   |
| History of COPD       | 6       | 4.11                        | 39 11.34                  | 37 11.6   |
| History of CVA/TIA    | 17      | 11.64                       | 41 11.92                  | 39 12.32  |
| Prior HF             | 35      | 23.97                       | 41 11.81                  | 39 12.6   |
| Prior MR             | 30      | 20.55                       | 10 3.09                   | 10 1.27   |
| History of PCI        | 41      | 28.08                       | 41 11.62                  | 39 12.32  |
| History of CABG       | 8       | 5.48                        | 17 4.94                   | 22 6.94   |
| History of Diabetes   | 40      | 27.48                       | 17 4.94                   | 22 6.94   |
| History of PVD        | 40      | 27.48                       | 17 4.94                   | 22 6.94   |
| Age (years)          | 86.43   | 57.53                       | 30 8.72                   | 27 8.46   |
| Weight (kg)          | 82.02   | 20.38                       | 24 6.98                   | 23 7.21   |
| Sodium (mEq/L)       | 137.17  | 3.85                        | 39 11.34                  | 37 11.6   |
| Potassium (mEq/L)    | 4.16    | 0.57                        | 41 11.92                  | 39 12.32  |
| HCO3 (mEq/L)         | 22.4    | 4.2                         | 17 9.29                   | 39 12.32  |
| BUN (mg/dL)          | 18.21   | 7.56                        | 17 9.29                   | 39 12.32  |
| Creatinine (mg/dL)   | 0.94    | 0.22                        | 39 12.32                  | 39 12.32  |
| WBC (10^3/mm^3)      | 12.36   | 5.54                        | 15 8.85                   | 15 8.85   |
| Hgb (g/dL)           | 13.21   | 2.37                        | 15 8.85                   | 15 8.85   |
| Hematocrit (%)       | 39.05   | 7.05                        | 39 8.85                   | 39 8.85   |
| Platelets (10^3/mm^3) | 218.85 | 70.53                       | 27 9.29                   | 27 9.29   |
| ALT (U/L)            | 73.57   | 121.42                      | 39 8.85                   | 39 8.85   |
| Total bilirubin (mg/dL) | 0.89 | 0.58                        | 39 8.85                   | 39 8.85   |
| INR                  | 1.24    | 0.29                        | 15 8.85                   | 15 8.85   |
| GFR (ml/min/1.73 m²) | 81.47   | 16.47                       | 81.93 14.31               | 80.95 15.37 |

Downloaded from http://ahajournals.org by on August 23, 2021
|                  | CSWG MI | DRR INDEPENDENTLY CLUSTERED | DRR WITH ASSIGNED CLUSTERS | CSWG HF |
|------------------|---------|-------------------------------|-----------------------------|---------|
|                  | I       | II                            | III                         | I       |
| Lactate (mEq/L)  | 3.45    | 3.15                          | 2.17                        | 7.89    | 4.21 |
| pH               | 7.30    | 7.35                          | 0.12                        | 7.21    | 0.16 |
| MAP (mmHg)       | 79.63   | 16.5                          | 74.98                       | 15.7    | 68.42 | 16.21 |
| DBP (mmHg)       | 64.71   | 15.46                         | 58.46                       | 14.47   | 57.25 | 16.91 |
| SBP (mmHg)       | 104.65  | 23.76                         | 103.57                      | 23.97   | 93.19 | 22.40 |
| CI (L/min/m²)    | 1.92    | 0.62                          | 1.93                        | 0.55    | 1.76  | 0.57  |
| CO (L/min)       | 3.72    | 1.3                           | 3.88                        | 1.67    | 3.45  | 1.41  |
| CPI (W/m²)       | 0.33    | 0.14                          | 0.32                        | 0.11    | 0.27  | 0.14  |
| LVEDD (mm)       | 8.88    | 0.93                          | 4.98                        | 0.95    | 4.45  | 0.77  |
| Heart rate (1/min)| 87.97  | 22.03                         | 89.84                       | 19.48   | 96.39 | 24.64 |
| PCWP (mmHg)      | 23.91   | 8.66                          | 25.78                       | 9.95    | 23.36 | 9.45  |
| PAPD (mmHg)      | 22.67   | 7.84                          | 24.82                       | 6.85    | 23.65 | 8.38  |
| PASP (mmHg)      | 42.6    | 13.9                          | 47.12                       | 13.53   | 41.57 | 13.86 |
| mean PAP (mmHg)  | 29.29   | 9.19                          | 32.32                       | 8.42    | 29.6  | 9.69  |
| RAP (mmHg)       | 12.98   | 5.77                          | 14.8                        | 7.13    | 16.81 | 6.48  |
| PAPI (arbitrary units) | 2.38  | 3.84                          | 1.82                        | 1.03    | 1.11  | 0.59  |
| RVSWI (mmHg * ml/m²) | 5.00  | 2.84                          | 5.33                        | 2.78    | 3.41  | 3.07  |

Note: The table above represents measured data for different parameters across various clusters in a study. The data includes lactate levels, pH, MAP, DBP, SBP, CI, CO, CPI, LVEDD, heart rate, PCWP, PAPD, PASP, mean PAP, RAP, PAPI, and RVSWI.
Figure S4: Sources of Each Cluster’s Baseline Differences from the Other Clusters

Chord Plots illustrate the association between clusters and clinically relevant groups of variables based on organ system function. A connection (or “chord”) from a cluster to a category signifies that at least one variable in this cluster was different from the other two clusters combined. Relative chord thickness corresponds to the relative influence of each organ system in determining the characteristics of each phenotype. Only non-imputed data was used for these graphs. Panel A shows all the phenotypes and panels B-D highlight each phenotype individually.
To assess external reproducibility of the derived clusters we applied consensus k-Means clustering to patients in the DRR validation cohort independently on the same variables. In these representative plots, each column represents one patient while each row displays the assigned clusters. Well-defined squares indicate stable clusters. These figures suggest stability of the clusters when 3 is picked as the number of clusters (k). DRR: Danish Retroshock Registry.
**Figure S6: Similarity of Clusters in the CSWG-MI Derivation and the DRR Validation Cohort**

Parallel coordinate plots comparing the tendencies of clinical parameters throughout the CSWG MI and DRR validation cohorts with respect to different clusters. A value of 1 signifies that the mean value for one cluster was one standard deviation higher than the mean value of the two clusters that are compared in the respective graph. The plots indicate how similar the properties of the clusters in CSWG-MI and DRR were, when clustered and analyzed separately. Furthermore, they reveal the resemblance of the clusters assigned to the validation cohort using the nearest centroid classifier (“DRR predicted”) as compared to the clusters from independent consensus K means clustering (“DRR clustered”). CSWG: Cardiogenic Shock Working Group; DRR: Danish Retroshock Registry.
Figure S7: Distribution of Patients in the Clusters by Cohort

Top panel: Pie charts depicting the percentage of patients per cluster in each cohort. For DRR, the results are shown twice: “DRR clustered” depicts the clusters of the patients when DRR was independently *de novo* clustered, while “DRR predicted” depicts the clusters when they were applied based on the centroids of the clusters in the CSWG MI derivation cohort. Patients in “DRR clustered” and “DRR predicted” were in the same cluster in 82% of the cases, indicating the similarity of the two methods. Bottom Panel: The t-distributed stochastic neighbor embedding (t-SNE) plots for a visual representation of the clustering in each cohort. For these plots, the 6 variables that were used for clustering were reduced to 2 variables which enables plotting the results in a two-dimensional graph.

CSWG: Cardiogenic Shock Working Group (Registry); DRR: Danish Retroshock Registry; MI: Myocardial Infarction; HF: Heart Failure
Figure S8: Phenotype Characteristics in the Validation Cohorts

Radar plots illustrate the specific characteristics of the cardiogenic shock phenotypes in the validation cohorts DRR and CSWG HF. Data was normalized across all phenotypes to a mean of 0 and an SD of 1. The dashed black line marks the mean (0), while every concentric gray line signifies a 0.1 SD difference from the overall mean. ALT: Alanine Aminotransferase; BUN: Blood urea nitrogen; CI: Cardiac Index; CO: Cardiac Output; CPI: Cardiac Power Index; CPO: Cardiac Power Output; CSWG HF: Cardiogenic Shock Working Group Registry Heart Failure Cohort; DBP: Diastolic Blood Pressure; DRR: Danish Retroshock Registry; GFR: Glomerular Filtration Rate; Hgb: Hemoglobin; INR: International Normalized Ratio; MAP: Mean Arterial Pressure; LVEDD: Left ventricular end-diastolic dimension; PA Sat: Pulmonary Arterial Saturation; PADP: Pulmonary Artery Diastolic Pressure; PAP: Pulmonary Artery Pressure; PAPI: Pulmonary Artery Pulsatility Index; PASP: Pulmonary Artery Systolic Pressure; PCWP: Pulmonary Capillary Wedge Pressure; RAP: Right atrial pressure; RVSWI: Right ventricular stroke work index; SBP: Systolic blood pressure; SVI: Stroke Volume Index WBC: White Blood Cell Count
Table S2: Sensitivity Analysis of DRR including all Patients

No patients were excluded for this analysis but missing values (overall missingness 4717 values = 8.1%) in DRR were imputed. The results resemble the results of the main analysis. ALT: Alanine Aminotransferase; BUN: Blood urea nitrogen; CABG: Coronary artery bypass graft; COPD: Chronic obstructive pulmonary disease; CVA/TIA: Cerebrovascular accident/Transient Ischemic Attack; DM2: Type 2 Diabetes Mellitus; DRR: Danish Retroshock Registry; ECMO: Extracorporeal membrane oxygenation; GFR: Glomerular Filtration Rate; Hgb: Hemoglobin; HTN: Hypertension; IABP: Intra-Aortic Balloon Pump; INR: International Normalized Ratio; MI: Myocardial infarction; PCI: Percutaneous coronary intervention; WBC: White Blood Cell Count.

| variable          | Non-congested | Cardiorenal | Cardiometabolic |
|-------------------|---------------|-------------|-----------------|
| n                 | 512           | 316         | 888             |
| Age               | 65.0 (12.0)   | 72.6 (10.1) | 66.9 (11.9)     |
| Gender            | 386 (75.4)    | 222 (70.3)  | 670 (75.5)      |
| Weight            | 78.2 (14.2)   | 78.6 (15.3) | 81.2 (14.9)     |
| DM2               | 59 (11.5)     | 64 (20.3)   | 150 (16.9)      |
| Prior MI          | 69 (13.5)     | 59 (18.7)   | 124 (14.0)      |
| HTN               | 242 (47.3)    | 216 (68.4)  | 440 (49.5)      |
| CVA/TIA           | 38 (7.4)      | 37 (11.7)   | 61 (6.9)        |
| COPD              | 50 (9.8)      | 40 (12.7)   | 87 (9.8)        |
| Prior PCI         | 483 (94.3)    | 300 (94.9)  | 842 (94.8)      |
| Prior CABG        | 38 (7.4)      | 16 (5.1)    | 47 (5.3)        |
| Pressors or Inotropes | 437 (85.4)    | 275 (87.0)  | 780 (87.8)      |
| IABP              | 64 (12.5)     | 40 (12.7)   | 84 (9.5)        |
| ECMO              | 13 (2.5)      |             | 45 (5.1)        |
| Mortality         | 183 (36.6)    | 155 (50.3)  | 558 (63.9)      |
| Lactate (mEq/L)   | 3.7 (2.6)     | 2.5 (1.3)   | 8.7 (4.6)       |
| HCO3 (mEq/L)      | 21.5 (3.8)    | 20.3 (3.5)  | 16.2 (4.2)      |
| GFR (mL/min/1.73 m²) | 80.2 (16.3)   | 35.8 (15.4) | 52.5 (18.7)     |
| Creatinine (mg/dL)| 0.9 (0.2)     | 2.3 (1.8)   | 1.5 (0.9)       |
| BUN (mg/dL)       | 18.4 (7.4)    | 36.4 (23.2) | 24.7 (15.2)     |
| Platelets (10^3/mm³)| 246.3 (86.5) | 240.5 (92.1)| 248.6 (87.8)    |
| Total bilirubin (mg/dL) | 0.7 (0.9)       | 0.7 (0.5)   | 0.7 (0.8)       |
| WBC (10^3/mm³)    | 14.1 (5.9)    | 15.1 (6.1)  | 17.2 (6.5)      |
| Hematocrit (%)    | 38.7 (6.5)    | 36.9 (6.4)  | 39.2 (7.2)      |
| Potassium (mEq/L) | 3.9 (0.6)     | 4.2 (0.8)   | 4.2 (0.9)       |
| Sodium (mEq/L)    | 137.7 (4.5)   | 137.3 (4.1) | 138.2 (4.6)     |
| Hemoglobin (g/dL) | 13.5 (1.9)    | 12.6 (2.0)  | 13.5 (2.2)      |
| ALT (U/L)         | 89.5 (132.1)  | 106.7 (225.3)| 255.5 (556.0)  |
| INR               | 1.2 (0.4)     | 1.3 (0.5)   | 1.3 (0.6)       |
Table S3: Sensitivity Analysis of CSWG HF including all Patients

No patients were excluded for this analysis but missing values (overall missingness 6161 values = 20.4%) in CSWG HF were imputed. The results resemble the results of the main analysis. ALT: Alanine Aminotransferase; BUN: Blood urea nitrogen; CABG: Coronary artery bypass graft; COPD: Chronic obstructive pulmonary disease; CSWG-HF: Cardiogenic Shock Working Group registry Heart Failure cohort; CVA/TIA: Cerebrovascular accident/Transient Ischemic Attack; DM2: Type 2 Diabetes Mellitus; ECMO: Extracorporeal membrane oxygenation; GFR: Glomerular Filtration Rate; Hgb: Hemoglobin; HTN: Hypertension; IABP: Intra-Aortic Balloon Pump; INR: International Normalized Ratio; MI: Myocardial infarction; PCI: Percutaneous coronary intervention; WBC: White Blood Cell Count.

| variable       | level          | Non-congested | Cardiorenal | Cardiometabolic |
|----------------|----------------|---------------|-------------|-----------------|
| n              |                | 227           | 379         | 80              |
| Age            | Years          | 52.1 (15.0)   | 60.5 (12.3) | 61.2 (15.0)     |
| Gender         | Male           | 169 (74.4)    | 300 (79.2)  | 52 (65.0)       |
| Weight         | kg             | 81.5 (19.7)   | 86.5 (19.8) | 88.4 (22.4)     |
| DM2            |                | 49 (21.6)     | 136 (35.9)  | 30 (37.5)       |
| Prior MI       |                | 52 (22.9)     | 128 (33.8)  | 25 (31.2)       |
| HTN            |                | 75 (33.0)     | 207 (54.6)  | 53 (66.2)       |
| CVA/TIA        |                | 28 (12.3)     | 49 (12.9)   | 7 (8.8)         |
| COPD           |                | 20 (8.8)      | 31 (8.2)    | 10 (12.5)       |
| Prior PCI      |                | 51 (22.5)     | 124 (32.7)  | 15 (18.8)       |
| Prior CABG     |                | 11 (4.8)      | 48 (12.7)   | 11 (13.8)       |
| Pressors or Inotropes |      | 177 (78.0)    | 319 (84.2)  | 66 (82.5)       |
| IABP           |                | 103 (45.4)    | 234 (61.7)  | 39 (48.8)       |
| ECMO           |                | 22 (9.7)      | 44 (11.6)   | 34 (42.5)       |
| Mortality      |                | 36 (16.4)     | 89 (24.3)   | 40 (50.6)       |
| Lactate        | mEq/L          | 2.3 (1.0)     | 2.1 (1.0)   | 9.2 (4.6)       |
| HCO3           | mEq/L          | 26.0 (2.9)    | 25.0 (2.9)  | 17.2 (4.0)      |
| GFR            | mL/min/1.73 m² | 76.1 (20.1)   | 33.6 (12.2) | 39.3 (17.7)     |
| Creatinine     | mg/dL          | 1.1 (0.2)     | 2.3 (1.2)   | 2.1 (1.2)       |
| BUN            | mg/dL          | 24.7 (9.3)    | 48.5 (20.6) | 37.6 (20.0)     |
| Platelets      | 10³/mm³        | 216.7 (66.9)  | 186.4 (60.2)| 200.1 (78.1)    |
| Total bilirubin| mg/dL          | 1.6 (1.9)     | 1.5 (1.3)   | 2.1 (1.5)       |
| WBC            | 10³/mm³        | 9.4 (3.8)     | 10.2 (4.2)  | 18.5 (13.7)     |
| Hematocrit     | %              | 37.7 (5.5)    | 35.9 (4.8)  | 36.2 (6.9)      |
| Potassium      | mEq/L          | 4.1 (0.5)     | 4.3 (0.5)   | 4.9 (0.9)       |
| Sodium         | mEq/L          | 134.8 (4.6)   | 133.3 (4.2) | 135.3 (5.4)     |
| Hemoglobin     | g/dL           | 12.4 (1.9)    | 11.8 (1.7)  | 11.8 (2.4)      |
| ALT            | U/L            | 82.9 (122.8)  | 163.4 (428.6)| 838.1 (1344.1) |
| INR            | 1.8 (0.9)      | 1.7 (0.9)     | 2.4 (1.6)   |
Figure S9: Sensitivity Analysis of DRR including all Patients

Patients in the Danish Retroshock Registry were excluded if missing any of the six values that were necessary for clustering (n=633). Depicted in this figure is the comparison of de novo consensus k means clustering and the cluster assignment based on the nearest centroids of the CSWG derivation cohort within the full (partially imputed) DRR cohort (n=1716). Cluster distribution differed between both methods, but the tendencies of the variables that were clustered on remained stable, underlining the external validity of the clusters.
Table S4. Odds of Mortality Associated with SCAI Stage and Phenotypes.

|       | CSWG MI OR (95% CI) | aOR* (95% CI) | CSWG HF OR (95% CI) | aOR* (95% CI) | CSWG MI+HF OR (95% CI) | aOR* (95% CI) |
|-------|---------------------|---------------|---------------------|---------------|------------------------|---------------|
| **SCAI** |                     |               |                     |               |                        |               |
| B     | 0 -                 | 0 -           | 0 -                 | 0 -           | 0 -                    | 0 -           |
| C     | 1.0 -               | 1.0 -         | 1.0 -               | 1.0 -         | 1.0 -                  | 1.0 -         |
| D     | 3.8 (1.9-7.7)       | 3.4 (1.7-6.9) | 6.1 (2.8-13.0)      | 5.6 (2.6-12.2) | 4.8 (2.8-8.0)          | 4.3 (2.6-7.2) |
| E     | 7.9 (3.7-16.8)      | 6.3 (2.9-13.6)| 14.2 (5.7-35.4)     | 9.3 (3.6-23.7)| 10.9 (6.1-19.4)        | 8.2 (4.5-14.7)|
| **Phenotypes** |                   |               |                     |               |                        |               |
| I     | 1.0 -               | 1.0 -         | 1.0 -               | 1.0 -         | 1.0 -                  | 1.0 -         |
| II    | 3.3 (2.3-4.8)       | 2.9 (1.7-4.9) | 3.2 (1.9-5.3)       | 2.9 (1.6-5.2) | 3.3 (2.3-4.8)          | 2.7 (1.9-4.0) |
| III   | 6.6 (4.3-10.0)      | 4.0 (2.3-7.1) | 4.7 (2.8-8.1)       | 6.1 (2.9-12.7)| 6.6 (4.3-10.0)         | 4.8 (3.0-7.5) |

*aORs adjusted by SCAI stage and CSWG derived clusters. aOR: adjusted odds ratio; CSWG: Cardiogenic Shock Working Group (Registry); HF: Heart Failure; MI: Myocardial Infarction; OR: odds ratio