Published in final edited form as:

J Clin Monit Comput. 2018 February ; 32(1): 117–126. doi:10.1007/s10877-017-0001-7.

Cardiorespiratory instability in monitored step-down unit patients: using cluster analysis to identify patterns of change

Eliezer L. Bose¹, Gilles Clermont², Lujie Chen³, Artur W. Dubrawski³, Dianxu Ren⁴, Leslie A. Hoffman⁴, Michael R. Pinsky², and Marilyn Hravnak⁴

¹The University of Texas at Austin, 1710 Red River St., Austin, TX 78701, USA

²Department of Critical Care Medicine, University of Pittsburgh Schools of Medicine, Pittsburgh, USA

³Auton Lab, Carnegie Mellon University Robotics Institute, Pittsburgh, USA

⁴Department of Acute and Tertiary Care, University of Pittsburgh Schools of Nursing, Pittsburgh, USA

Abstract

Cardiorespiratory instability (CRI) in monitored step-down unit (SDU) patients has a variety of etiologies, and likely manifests in patterns of vital signs (VS) changes. We explored use of clustering techniques to identify patterns in the initial CRI epoch (CRI₁; first exceedances of VS beyond stability thresholds after SDU admission) of unstable patients, and inter-cluster differences in admission characteristics and outcomes. Continuous noninvasive monitoring of heart rate (HR), respiratory rate (RR), and pulse oximetry (SpO₂) were sampled at 1/20 Hz. We identified CRI₁ in 165 patients, employed hierarchical and *k*-means clustering, tested several clustering solutions, used 10-fold cross validation to establish the best solution and assessed inter-cluster differences in admission characteristics and outcomes. Three clusters (C) were derived: C1) normal/high HR and RR, normal SpO₂ (n = 30); C2) normal HR and RR, low SpO₂ (n = 103); and C3) low/normal HR, low RR and normal SpO₂ (n = 32). Clusters were significantly different based on age ($p < 0.001$; older patients in C2), number of comorbidities ($p = 0.008$; more C2 patients had 2) and hospital length of stay ($p = 0.006$; C1 patients stayed longer). There were no between-cluster differences in SDU length of stay, or mortality. Three different clusters of VS presentations for CRI₁ were identified. Clusters varied on age, number of comorbidities and hospital length of stay. Future study is needed to determine if there are common physiologic underpinnings of VS clusters which might inform clinical decision-making when CRI first manifests.

Correspondence to: Eliezer L. Bose.

Compliance with ethical standards

Ethical approval

The study was approved and has current active approval from the University of Pittsburgh Institutional Review Board (PRO12070002). All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. For this type of study formal consent is not required, and was approved with consent waiver.

Conflict of interest The authors have no commercial conflicts of interest to report.

Keywords

Cardiorespiratory instability; Cluster analysis; Non-invasive monitoring

1 Introduction

Patients are placed in step-down units (SDUs) because their presumed risk for cardiorespiratory instability (CRI) necessitates a higher level of monitoring than is available in general hospital units [1]. Although multiple vital signs (VS) are continuously monitored, alerts are provided to staff only when individual parameters exceed presumed thresholds of stability, without relationship to each other. To better inform clinical decision-making, it would be helpful to know if VS changes manifest in unique clusters, and are related to certain patient characteristics or pathophysiologic etiologies.

Initial exploration of the potential of grouping patient deterioration according to VS change patterns was motivated by the emergence of medical emergency teams (MET) [2], prompting evaluation of their benefit to improve patient outcomes [2–6], and critical analysis of events preceding a MET call for help [7–9]. Jones et al. [8] studied 400 MET calls in intensive care unit (ICU) settings and found the underlying reasons for initiating calls were hypoxia (41%), hypotension (28%), altered consciousness (23%), tachycardia (19%), tachypnea (14%) and oliguria. They termed these etiologies “MET syndromes.” In their study, infections (especially pneumonia), cardiogenic shock or pulmonary edema and arrhythmias were responsible for 53% of all MET calls. Jones et al. were the first to suggest that, since there are differing causes for MET calls, there must be differing patterns in monitored variables that precede the unstable state. If so, they suggested there might be different approaches to managing CRI based on MET syndrome. We were unable to identify any studies that attempted to analyze patterns of CRI using commonly monitored variables, e.g., heart rate (HR), respiratory rate (RR) and peripheral oxygen saturation (SpO₂) prior to instability that was sufficiently severe to prompt MET activation. Having such information prior to a MET need might help clinicians recognize the syndrome, or instability pattern, earlier in its course and provide more targeted supportive or preventative measures to avert the crisis.

The subtle and interactive changes which occur in early instability are complex and multiple, making it difficult to detect patterns of change. Machine learning (ML) [10] entails using unsupervised learning algorithms in an attempt to find subtle and unrecognized patterns within physiologic data [11–13]. Advanced computational algorithms can group subjects into clusters which share similar properties. The most common method to accomplish this goal involves hierarchical and *k*-means clustering [14] to find hidden patterns, or groupings. Monitored variables are grouped using metrics for similarity such as squared Euclidean distance [15] to link subjects within a cluster. These clustering techniques are used to find patterns in bioinformatics for sequence analysis and genetic clustering [16], in data mining [17], and medical imaging [18].

The purpose of this study was to explore use of clustering techniques to provide insight into VS patterns of SDU patients at the first manifestation of CRI and determine if there were inter-cluster differences in admission characteristics or outcomes.

2 Materials and methods

2.1 Patients and setting

Following Institutional Review Board approval, we collected continuous VS data streams and hospital admission characteristics from 634 patient admissions to a 24-bed adult SDU of a Level-1 Trauma Center during an 8-week period from November 16, 2006 to January 11, 2007. Continuous VS data streams included HR (3-lead ECG), RR (bio impedance) and SpO₂ (pulse oximeter Model M1191B, Phillips, Boeblingen, Germany) recorded at 1/20 Hz frequency.

2.2 Event identification (CRI₁ determination)

VS excursions beyond our local criteria for instability concern (HR < 40 or > 140, RR < 8 or > 36, SpO₂ < 85%) not due to artifact were defined as CRI if they persisted for a minimum duration of 3 min, or for 4 min out of a 5-minute moving window (80% duty cycle) with or without changes in other VS. Of the 634 patients, 195 (31%) experienced CRI and 439 (69%) did not. For this paper, analysis of data was restricted to the first CRI event (CRI₁). The first vital sign to exceed the previously identified thresholds was termed the CRI₁ driver.

2.3 Cluster identification

Steps in preprocessing were feature extraction, standardization, and outlier removal. A combination plot of the feature space of all variables was created to facilitate visualization of all possible combinations between VS for each of the features. Hierarchical and *k*-means clustering were then used to complete analysis of the data.

2.4 Feature extraction, standardization and outlier removal

We extracted statistical features (mean, median, mode, minimum, maximum, range, variance and standard error of mean over CRI₁ duration) for HR, RR, and SpO₂. Since measurement of each VS used a different scale, we standardized the variables, thus rendering them scale free. Clustering techniques and *k*-means in particular are sensitive to outliers, causing outliers to be selected as initial clusters. Two patients identified in this way were eliminated ($|z| > 3.29$), leaving 193 patients.

2.5 Combination plot of feature space

We developed a combination plot of the feature space to evaluate for missingness among the features. In this plot, if a feature for HR exists but RR does not exist, it is color coded, indicating missingness. We discovered 13 patients with only HR and RR, 7 patients with only HR and SpO₂, 5 patients with missing data for HR or SpO₂ and 3 patients with only SpO₂. These 28 patients were omitted, leaving 165 patients with complete data for cluster analysis.

2.6 Hierarchical clustering

We next implemented agglomerative hierarchical clustering on the remaining sample. In order to form clusters, there must be a criterion for similarity or distance between patients and a criterion for determining which clusters are to be linked at successive steps. We used squared Euclidean distance, since all variables were continuous [19], and Ward's linkage method for determining linkage of clusters [20]. Clustering began with each case being a cluster unto itself. The squared Euclidean distance to cluster means was calculated and distances summed for all cases. Next, cases with smaller distances from each other were linked until all clusters were completed and means for all variables were calculated for each cluster. At each stage, the two clusters that merged were those that resulted in the smallest increase in the overall sum of squares within-cluster distances [21, 22]. A visual representation of the distance at which clusters combined was represented by a dendrogram.

2.7 K-means clustering

The major difference in *k*-means involves pre-specifying *k*, which is the number of clusters. By visualizing potential cluster solutions from the dendrogram, we chose an initial starting value for *k* at 3. The first step in *k*-means clustering involved finding *k* centers. The algorithm started with an initial set of means and classified patients based on their distances to the centers. Next, it re-computed the cluster means using patients that were assigned to the cluster and then reclassified all patients based on a new set of means. The steps were then repeated until cluster means changed minimally between successive steps. Finally, the algorithm calculated the means of the clusters again and assigned patients to their final clusters [23, 24]. We tested several clustering solutions that entailed using 3 to 5 clusters after visualizing the dendrogram obtained from hierarchical clustering.

2.8 Feature selection

Following initial identification of clusters, an ANOVA table was constructed to describe F-ratios per feature within the clusters for a particular solution. If non-significance was observed, it indicated that this feature contributed minimally to separation of clusters. From this analysis, we discovered that VS range, variance and standard error of mean value were mostly non-significant whereas features of mean, median, mode, maximum and minimum always contributed significantly to differences between clusters. Hence, we used the latter 5 features per VS for further clustering analyses. However, upon finding that all of these derived features were highly correlated ($r > 0.9$, $p < 0.001$), we used median alone as a feature for clustering analysis.

2.9 Cross validation of cluster solution

As the final step, we implemented cross-validation (CV) of each of the cluster solutions. To determine the best solution, we performed a commonly used process [25] termed a stratified 10-fold CV to each solution. For this analysis, the original sample was partitioned into 10 equal sized sub-samples. A single subsample was retained as validation set for testing the cluster solution with the remaining 9 sub-samples used as training data and the process repeated 10 times. The advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once.

2.10 CRI₁ driver within each cluster of patients

Once we identified the best cluster solution, we labeled each cluster based on the general pattern of each VS within that particular cluster. We also examined the interaction among VS within that cluster and its relationship with the other two VS per patient, using cross correlations of HR with RR, HR with SpO₂ and RR with SpO₂ per patient within each cluster.

2.11 Admission characteristics and outcomes

Admission characteristics were age, gender, race, type of service (admitted to a medical or surgical service, irrespective of whether they had a later surgery or not), medical history, and Charlson-Deyo comorbidity Index (0, 1, 2) [26]. We recorded admission diagnosis based on the ICD-9 classification code assigned at hospital admission, which we first grouped into 14 categories according to the classification system reported by Brown et al. [27], and further reduced into 4 categories based on code frequency. The resulting 4 diagnosis categories were: trauma, diseases of circulatory system, diseases of digestive system, and all others. We also recorded SDU admission source (transfer from a higher intensity monitoring unit, or direct SDU admission or transfer from lower intensity unit). Outcomes were hospital length of stay (HLOS) in days, SDU length of stay (SDU LOS), and discharge disposition from the SDU (transfer to a higher or lower intensity unit). Lower intensity unit collectively included patients transferred either to a rehabilitation center, a long-term hospital, a skilled nursing facility, a home health agency or directly to home. In our sample, there were no reported deaths.

2.12 Statistical analysis

Data were analyzed using IBM SPSS v22 for descriptive analysis and exploration, WEKA and MATLAB R2015a for clustering and cross validation. One-way ANOVA was used to analyze cluster differences between groups for continuous variables and Chi square to compare categorical variables. A $p < 0.05$ was considered significant.

3 Results

Of 634 patients, 195 (31%) developed CRI₁ due to HR, RR or SpO₂ and 165 (85%) were used for feature extraction. A summary of features extracted from CRI₁ epochs per VS is provided in Table 1.

3.1 Hierarchical and k-means clustering

Results of hierarchical clustering and distribution of the 165 patients into clusters (C) are shown in Fig. 1. For a 3-cluster solution, results were: C1 62%, C2 19% and C3 19%. Using this information, we tested *k*-means clustering with *k*-values set between 3 and 5. Since *k*-means uses a different algorithm compared to hierarchical clustering, for $k = 3$ the results were: C1 18%, C2 62%, C3 19% (Fig. 2). With $k = 4$ the results were: C1 23%, C2 41%, C3 19% and C4 16%.

3.2 Cross validation

For the 3 cluster solution, 99.4% of instances were correctly classified (Table 2). The sensitivity and specificity values were C1 (96.7%, 100%), C2 (100%, 98.4%) and C3 (100%, 100%). For C1, *k*-means was less sensitive but more specific. For C2 and C3, *k*-means clustering correctly identified instances belonging to each of the clusters respectively. For the 4 cluster solution, 89.7% of instances were correctly classified (Table 2). The sensitivity and specificity values were C1 (60.5%, 98.4%), C2 (98.5%, 84.5%), C3 (100%, 100%) and C4 (96.3%, 100%). The percentage of patients belonging to C3 alone were accurately identified, but was less sensitive and more specific with the other clusters. For the 5 cluster solution, *k*-means performed much poorly achieving only 80% accuracy (Table 2). However, overall, a 3-cluster solution performed better than a 4-cluster or 5-cluster solution due to a higher percentage of accurately classified instances.

3.3 VS groupings in 3-cluster solution

As shown in Fig. 2, in C1, there were 30 patients whose VS grouped into normal/high (defined as within the parameter thresholds for instability concern previously described) HR (70–164 bpm), normal/high RR (17–44 breaths/min) and normal SpO₂ (85–100%) with the driver being primarily either RR or HR. In C2 there were 103 patients, whose VS grouped into normal HR (55–140 bpm), normal RR (9–32 breaths/min), and low SpO₂ (73–85%). Almost every patient within C2 had SpO₂ as the driver. In C3, there were 32 patients, and VS grouped as low/normal HR (35–116 bpm), low RR (5–13 breaths/min), and normal SpO₂ (89–100%), with the driver for almost all patients being RR. Significant cross-correlations ($p < 0.05$) were observed only between HR with RR. In C2 (the majority group), HR and RR had a positive cross-correlation, either increasing or decreasing simultaneously. For C3, HR and RR had a negative cross-correlation indicating that while that HR increased RR decreased, and while RR increased, HR decreased.

3.4 Admission characteristics and outcomes per cluster

There were significant differences between clusters in age and Charlson Deyo comorbidity index (Table 3). C2 contained mostly older patients (age $\mu = 65.8$, $\sigma = 18$). C2 had a higher proportion (34%) of patients with higher Charlson Comorbidity Index (≥ 2), indicating that patients within this cluster also had more comorbid conditions upon admission. The outcome of hospital LOS was significant (Table 4). C1 patients stayed the longest (LOS $\mu = 25.2$, $\sigma = 47$). However, SDU LOS did not differ between patients by cluster.

4 Discussion

This study is the first to use hierarchical and *k*-means clustering on the feature space of CRI₁ epoch to derive the VS pattern clusters of CRI₁ in SDU patients. There are four key findings. First, the most common driver of CRI₁ in this population was SpO₂. Second, three main clusters were derived, and third, these clusters contained predictable patterns for drivers and accompanying VS behavior: C1) normal/high HR and RR but normal SpO₂; C2) normal/HR and RR and low SpO₂; C3) low/normal HR, low RR and normal SpO₂. Finally, admission characteristics between patients within each cluster differed in regard to age, and Charlson

Deyo comorbidity index with the outcome of hospital LOS significantly different between the clusters.

Using only median values for the VS of HR, RR and SpO₂ for CRI₁ epoch and applying clustering techniques, we found that a 3-cluster solution was best. An interesting finding was that within two of the clusters, the driver was identical for almost all patients in that cluster. For C2, the driver was SpO₂, for C3 RR. In C1 (normal/high HR and RR but normal SpO₂), there was only 1 patient whose driver was SpO₂, while for the rest it was either HR or RR. On review of the actual data within the CRI₁ epoch for this patient, we discovered that this patients' RR was 32–36 breaths/min, (i.e. relatively high) similar to other patients in the cluster. However, due to our abnormal threshold definition (RR > 36), this patients' CRI₁ driver was considered not to be RR but instead low SpO₂ < 85%.

We initially decided to use a total of 24 features derived from the CRI epoch per VS, which included the features shown in Table 1, similar to our previous work [28]. However, in testing clustering techniques, the only features which were consistently found to contribute to clustering were the mean, median, mode, minimum and maximum value. A commonly employed procedure is to use ANOVA similar to our approach and only retain features which contributed significantly towards clustering [29, 30]. Besides this, we also observed that Q-Q plots generated for the features of range, variance and standard error of the mean per VS showed fat tails and severe departures from normality for the distribution. Further, extremely high kurtosis values were observed especially for HR values. Removing outlier cases would have resulted in further loss of patients available for clustering. We chose to eliminate [31, 32] these 9 features instead of the patients, leaving 15 features for cluster analysis. Upon further discovering that there were correlations among the features, we finally used 3 features (median values of HR, RR and SpO₂) for cluster analysis.

We employed the most commonly used unsupervised machine learning techniques of hierarchical and *k*-means clustering to derive CRI patterns [33] and used them in a complementary way, with each approach giving slightly different information. We used hierarchical clustering and visualization of the dendrogram to potentially estimate where *k*-value could be set for *k*-means clustering. Our choice of a bottom-up approach such as agglomerative hierarchical clustering was to ensure that nearby and thus similar cases ended up within the same cluster, as opposed to top-down approaches [34]. While heuristic techniques exist for deciding on an optimal value for the number of clusters such as the use of an elbow plot [35] or other techniques [15, 36], our use of hierarchical clustering was primarily to estimate potential *k*-values, similar to others [30, 37]. Our choice of Ward's linkage method was that besides being a commonly used approach [38], it used ANOVA to evaluate distances between clusters in an attempt to minimize sum of squares of any two clusters [20], thus a more rigorous approach. Further, since our goal was to use *k*-means clustering also, Ward's linkage method was ideal for comparing the two procedures. This is because it is also based on the same analysis of the clustering problem as *k*-means and usually gives clusters that appear well-defined [39]. After visualization of the dendrogram, we decided on limiting the *k*-value used in *k*-means clustering to between 3 and 5 clusters.

We implemented 10-fold stratified cross validation (CV) [40] to determine if a 3-cluster or 4-cluster solution was better. Given our small dataset, stratified cross-validation was used so that all the folds would likely contain at least a good proportion of patients even from the smallest cluster. We evaluated the performance of k -means algorithm by its accuracy, which is the degree of success in classifying correctly new instances as shown in Table 2. While there is no universal algorithm giving the best performance in all possible learning situations [41], CV was particularly suited in our study given the narrowed down option of two possible cluster solutions, as opposed to other boot-strapping procedures [42] to decide the best cluster solution.

We used cross-correlations among the VS within the cluster to study the interactions within the 3 VS. For instance, a negative cross correlation between RR and SpO₂ in a particular case would indicate that RR and SpO₂ were moving in the opposite direction while a positive cross-correlation would suggest a parallel effect. We expected that all patients within the cluster would exhibit certain pattern of changes for the three VS, but this was not noticed. This suggests that while there are broad cluster patterns within CRI epoch, there could be subsets of patterns within the cluster not captured by our analysis or subtle variations within VS even prior to entering into CRI epoch [43, 44]. Future studies are needed to explore this potential.

The only admission characteristics that differed significantly between the patients in each cluster were age and comorbidity Index. In the largest cluster (C2), mean age was 65.8 and the driver for CRI₁ was SpO₂, indicating that older patients were more likely to develop CRI₁ due to SpO₂, with the pattern being normal HR, normal RR and low SpO₂. From a clinical perspective, this indicates that older patients with increased comorbid conditions should be monitored more closely for changes in SpO₂.

Finally, we chose to consider only the first instance of CRI (CRI₁) in this study. The clinical utility of our study is to aid CRI pattern recognition so as to eventually target surveillance and intervention based on probable cluster patterns, right from the first occurrence of CRI. The clusters we obtained are limited by being descriptive. However, this study is a stepping stone to identify if patients continue to remain within the same cluster, or if they eventually switch to a different pattern, should successive instances of CRI occur. Future studies to track cluster progression will explore whether or not patients change clusters during further instances of CRI.

5 Limitations

VS analyzed for this study were obtained from the SDU of one Level-1 Trauma Center and may not be reflective of other hospitals or patient populations. As such, while we employed an interesting methodology, our results are limited. The greatest limitation to our study was a low number of subjects in the study cohort. It is possible that with a greater number of patients, more clusters with less overlap might have been discovered which would have been more informative for both research and clinical application. It is also possible that enriching the data pool by expanding the data elements under consideration beyond basic demographics and array of data features we used might have resulted in more clinically

relevant clusters. The parameter thresholds used contain normality, although commonly used as cut points for serious instability [1, 45, 46], are in fact beyond normal values for a stable physiologic state. Use of lower thresholds of lesser instability magnitude might have altered our results. Also, we recorded our data at 1/20 Hz for each of the VS, and within this recording frequency, there were a few instances when all values were not recorded. Nevertheless, we did not use missing value imputation or any other procedures commonly employed in time-series cross-section data to account for missing data [47, 48]. We made this decision to insure the data were as accurate and as close to normal SDU monitoring practices as possible, yet seek features we could use for clustering. Finally, we employed a limited set of features, and results could differ with increased features derived from CRI₁ epoch, although the parsimonious feature set we used enabled testing both clustering approaches with less than 10 iterations of the algorithms.

6 Conclusion

Our analysis, using clustering techniques on the feature space of the VS of HR, RR and SpO₂ on CRI₁ epochs in SDU patients, derived 3 different clusters of VS presentations. The most common driver of CRI₁ was SpO₂ with the cluster pattern of normal HR and RR and low SpO₂. The cluster groups varied on age, number of comorbidities and hospital length of stay. Future study will be needed to determine if there are common physiologic underpinnings of the VS clusters which might inform clinical decision-making even at the point where CRI first manifests.

Acknowledgments

This study was funded with grant support from the United States National Institutes of Health, National Institute of Nursing Research RO1 NR13912 and the National Science Foundation NSF 1320347.

Funding

This study was funded with grant support from the United States National Institutes of Health, National Institute of Nursing Research RO1 NR13912 and the National Science Foundation NSF 1,320,347. The funding bodies approved the study design as submitted in the grant application proposal, but had no role in the data collection, analyses, or interpretation, manuscript preparation, or decision to submit the manuscript for publication.

References

1. Hravnak M, Edwards L, Clontz A, Valenta C, DeVita MA, Pinsky MR. Defining the incidence of cardiorespiratory instability in patients in step-down units using an electronic integrated monitoring system. *Arch Intern Med*. 2008; 168(12):1300–8. [PubMed: 18574087]
2. Buist MD, Moore GE, Bernard SA, Waxman BP, Anderson JN, Nguyen TV. Effects of a medical emergency team on reduction of incidence of and mortality from unexpected cardiac arrests in hospital: preliminary study. *Bmj*. 2002; 324(7334):387–90. [PubMed: 11850367]
3. Hillman K, Chen J, Cretikos M, Bellomo R, Brown D, Doig G, Finfer S, Flabouris A, Investigators MS. Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. *The Lancet*. 2005; 365(9477):2091–7.
4. Bellomo R, Goldsmith D, Uchino S, Buckmaster J, Hart GK, Opdam H, Silvester W, Doolan L, Gutteridge G. A prospective before-and-after trial of a medical emergency team. *Med J Aust*. 2003; 179(6):283–8. [PubMed: 12964909]
5. Bristow PJ, Hillman KM, Chey T, Daffurn K, Jacques TC, Norman SL, Bishop GF, Simmons EG. Rates of in-hospital arrests, deaths and intensive kcare admissions: the effect of a medical emergency team. *Med J Aust*. 2000; 173(5):236–40. [PubMed: 11130346]

6. DeVita M, Braithwaite R, Mahidhara R, Stuart S, Foraida M, Simmons R. Use of medical emergency team responses to reduce hospital cardiopulmonary arrests. *Quality Safety Health Care*. 2004; 13(4):251–254.
7. Downey AW, Quach JL, Haase M, Haase-Fielitz A, Jones D, Bellomo R. Characteristics and outcomes of patients receiving a medical emergency team review for acute change in conscious state or arrhythmias*. *Crit Care Med*. 2008; 36(2):477–81. [PubMed: 18091535]
8. Jones D, Duke G, Green J, Briedis J, Bellomo R, Casamento A, Kattula A, Way M. Medical emergency team syndromes and an approach to their management. *Critical care*. 2006; 10(1):R30. [PubMed: 16507153]
9. Jones D, Bates S, Warrillow S, Goldsmith D, Kattula A, Way M, Gutteridge G, Buckmaster J, Bellomo R. Effect of an education programme on the utilization of a medical emergency team in a teaching hospital. *Intern Med J*. 2006; 36(4):231–6. [PubMed: 16640740]
10. Alpaydin, E. Introduction to machine learning. 2nd. Boston: MIT press; 2014.
11. Jang E-H, Park B-J, Kim S-H, Sohn JH. Emotion recognition by machine learning algorithms using psychophysiological signals. *Int J Eng Ind*. 2012; 3(1)
12. AlZoubi O, Mello SK, Calvo R. Detecting naturalistic expressions of nonbasic affect using physiological signals. *IEEE Trans Affect Comput*. 2012; 3(3):298–310.
13. Martinez HP, Bengio Y, Yannakakis GN. Learning deep physiological models of affect. *IEEE Comput Intell Mag*. 2013; 8(2):20–33.
14. Jain AK, Murty MN, Flynn PJ. Data clustering: a review. *ACM computing surv (CSUR)*. 1999; 31(3):264–323.
15. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *J R Stat Soc B*. 2001; 63(2):411–23.
16. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences*. 1998; 95(25):14863–8.
17. Agrawal, R. Data Engineering, 1995. IEEE; Srikant R Mining sequential patterns; p. 3-14. Proceedings of the Eleventh International Conference on, 1995
18. Larrañaga P, Calvo B, Santana R, Bielza C, Galdiano J, Inza I, Lozano JA, Armañanzas R, Santafé G, Pérez A. Machine learning in bioinformatics. *Brief Bioinform*. 2006; 7(1):86–112. [PubMed: 16761367]
19. Sarstedt, M., Mooi, E. A concise guide to market research: the process, data, and methods using IBM SPSS statistics. Berlin: Springer; 2014.
20. Ward JH Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*. 1963; 58(301):236–44.
21. Jain, AK., Dubes, RC. Algorithms for clustering data. New Jersey: Prentice-Hall, Inc; 1988.
22. Murtagh F. A survey of recent advances in hierarchical clustering algorithms. *The Computer Journal*. 1983; 26(4):354–9.
23. Steinley D. K-means clustering: a half-century synthesis. *Br J Math Stat Psychol*. 2006; 59(1):1–34. [PubMed: 16709277]
24. Jain AK. Data clustering: 50 years beyond K-means. *Pattern recognition letters*. 2010; 31(8):651–66.
25. McLachlan, G., Do, K., Ambrose, C. Analyzing microarray gene expression data. Vol. 422. New Jersey: John Wiley & Sons; 2005.
26. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol*. 1992; 45(6):613–9. [PubMed: 1607900]
27. Brown H, Terrence J, Vasquez P, Bates DW, Zimlichman E. Continuous monitoring in an inpatient medical-surgical unit: a controlled clinical trial. *Am J Med*. 2014; 127(3):226–32. [PubMed: 24342543]
28. Hravnak M, Chen L, Dubrawski A, Clermont G, Bose E, Fiterau M, Wang D, Guillame-Bert M, Pinsky M. Supervised Machine Learning Can Classify Artifact In Multi-Signal Vital Sign Monitoring Data From Step-Down Unit (SDU) Patients. *Intensive Care Medicine*. 2014:S28.

29. Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, Wardlaw AJ, Green RH. Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med*. 2008; 178(3):218–24. [PubMed: 18480428]
30. Cellucci T, Tyrrell PN, Twilt M, Sheikh S, Benseler SM. Distinct phenotype clusters in childhood inflammatory brain diseases: implications for diagnostic evaluation. *Arthritis & Rheumatology*. 2014; 66(3):750–6.
31. Devaney M, Ram A. Efficient feature selection in conceptual clustering. *ICML*. 1997:92–97.
32. Talavera L. Feature selection as a preprocessing step for hierarchical clustering. *International Conference of Machine Learning Proceedings; ICML*. 1999; p. 389–397.
33. Hastie, T., Tibshirani, R., Friedman, J. *The Elements of Statistical Learning*. Second. New York: Springer; 2009. Unsupervised learning; p. 485–585.
34. Arai K, Barakbah AR. Hierarchical K-means: an algorithm for centroids initialization for K-means. *Reports of the Faculty of Science and Engineering*. 2007; 36(1):25–31.
35. Mooi, E., Sarstedt, M. *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*. Berlin: Springer-Verlag; 2011. Cluster analysis; p. 273–98.
36. Milligan GW, Cooper MC. An examination of procedures for determining the number of clusters in a data set. *Psychometrika*. 1985; 50(2):159–79.
37. Hawkins MA, Schaefer JT, Gunstad J, Dolansky MA, Redle JD, Josephson R, Moore SM, Hughes JW. What is your patient's cognitive profile? Three distinct subgroups of cognitive function in persons with heart failure. *Appl Nurs Res*. 2015; 28(2):186–91. [PubMed: 25510559]
38. Stuetzle W, Nugent R. A generalized single linkage method for estimating the cluster tree of a density. *Journal of Computational Graphical Statistics*. 2012; 19(2):397–418.
39. Clarke, B., Fokoue, E., Zhang, HH. *Principles and theory for data mining and machine learning*. New York: Springer Science & Business Media; 2009.
40. Wagstaff K, Cardie C, Rogers S, Schrödl S. Constrained k-means clustering with background knowledge. *International Conference on Machine Learning (ICML)*. 2001:577–584.
41. Schaffer C. A conservation law for generalization performance. *Proceedings of the 11th international conference on machine learning*. 1994:259–265.
42. Duda, RO., Hart, PE., Stork, DG. *Pattern classification*. 2nd. New York: Wiley; 2012.
43. Clermont G, Sileanu F, Pinsky M, Ogundele O. Physiologic Variable Analysis Between Hemodynamic Stable And Unstable Patients. *Am J Respir Crit Care Med*. 2012; 185:A4630.
44. Clermont G, Sileanu F, Pinsky M, Ogundele O. Use Of Derived Physiologic Variables To Predict Individual Patients' Probability Of Hemodynamic Instability. *Am J Respir Crit Care Med*. 2013; 187:A5067.
45. Smith GB, Prytherch DR, Schmidt PE, Featherstone PI, Higgins B. A review, and performance evaluation, of single-parameter “track and trigger” systems. *Resuscitation*. 2008; 79(1):11–21. [PubMed: 18620794]
46. Hravnak M, DeVita MA, Clontz A, Edwards L, Valenta C, Pinsky MR. Cardiorespiratory instability before and after implementing an integrated monitoring system. *Crit Care Med*. 2011; 39(1):65. [PubMed: 20935559]
47. Honaker J, King G. What to do about missing values in time-series cross-section data. *American Journal of Political Science*. 2010; 54(2):561–81.
48. Honaker J, King G, Blackwell M. *Amelia II: A program for missing data*. *Journal of Statistical Software*. 2011; 45(7):1–47.

Fig. 1.

Dendrogram showing cluster distances. The dendrogram is read from bottom to top, with the vertical lines showing joined clusters. The bottom most row contains all $N = 165$ patients. Hierarchical clustering and Ward linkage method joined the cases. Panel B shows the cut dendrogram to obtain a 3-cluster solution. The majority (61%) of cases were in cluster 1. Cluster distances are plotted with the dendrogram which is read from bottom to top with the vertical lines showing joined clusters. The plot is rescaled to fall within the range of 1 to 25; however the ratio of the rescaled distances within the dendrogram is the same as the ratio of the original distances

Fig. 2.

K-Means cluster solution. For a 3-cluster solution, with $k = 3$, there were 30 patients in cluster 1 (C1), 103 patients in cluster 2 (C2) and 32 patients in cluster 3 (C3). The axis contains the standardized values of the median for heart rate (HR), respiratory rate (RR) and arterial pulse oximetry (SpO₂). The largest cluster is on the bottom of the 3D-graph. The pattern for this group is normal HR and RR with low SpO₂. A summary of the combination of the 3 vital signs per cluster is shown in the table below

Table 1

Numeric features extracted from initial cardiorespiratory instability (CRI₁) epoch by vital signs of heart rate (HR), respiratory rate (RR) and arterial pulse oxygenation (SpO₂) in N = 165 patients

Feature	HR Mean $\pm$ SD	RR Mean $\pm$ SD	SpO ₂ Mean $\pm$ SD
Mean	91 $\pm$ 23	19 $\pm$ 9	87 $\pm$ 8
Median	91 $\pm$ 23	19 $\pm$ 9	87 $\pm$ 7
Mode	90 $\pm$ 23	19 $\pm$ 9	87 $\pm$ 8
Maximum	97 $\pm$ 25	23 $\pm$ 10	90 $\pm$ 6
Minimum	86 $\pm$ 22	15 $\pm$ 7	84 $\pm$ 10

Table 2

Stratified cross-validation summary

Column name	3-Cluster solution			4-Cluster solution			5-Cluster solution		
	Number of Instances	(%)		Number of Instances	(%)		Number of Instances	(%)	
Correctly Classified Instances	164	99.4		148	89.7		132	80	
Incorrectly Classified Instances	1	0.6		17	10.3		24	14.5	
Unclassified Instances	0	0		0	0		9	5.5	
Cluster Number	Sensitivity (%)	Specificity (%)		Sensitivity (%)	Specificity (%)		Sensitivity (%)	Specificity (%)	
1	96.7	100		60.5	98.4		73.7	98.5	
2	100	98.4		98.5	84.5		93.7	84.9	
3	100	100		100	100		100	100	
4				96.3	100		95.5	100	
5							39.1	94	

All 3 cluster solutions achieved 100% sensitivity and specificity on at least 2 clusters. However, in terms of overall performance, a 3-cluster solution correctly classified more instances (99.4%), than a 4-cluster solution (89.7%) or a 5-cluster solution (80%)

Admission characteristics in patients with first instance of cardiorespiratory instability (CRI₁) event during the course of their stay on the step-down unit (SDU) by cluster

Table 3

Variable	Total (N = 165) No. (%)	Cluster 1 (n = 30) No. (%)	Cluster 2 (n = 103) No. (%)	Cluster 3 (n = 32) No. (%)	p-value
Age (Mean ± SD)	59.9 ± 21	51.4 ± 25	65.8 ± 18	49.1 ± 19	< 0.001 [*]
Gender					0.619
Male	85 (51.5)	16 (53.3)	55 (53.4)	14 (43.8)	
Female	80 (48.5)	14 (46.7)	48 (46.6)	18 (56.3)	
Race					0.970
White	129 (78.2)	23 (76.7)	82 (79.6)	24 (75.0)	
Black	17 (10.3)	3 (10.0)	10 (9.7)	4 (12.5)	
Other	19 (11.5)	4 (13.3)	11 (10.7)	4 (12.5)	
Type of Service					0.633
Surgical	86 (52.1)	18 (60.0)	52 (50.5)	16 (50.0)	
Medical	79 (47.9)	12 (40.0)	51 (49.5)	16 (50.0)	
Hospital Admission Diagnosis					0.394
Trauma	85 (51.5)	18 (60.0)	54 (52.4)	13 (40.6)	
Diseases of Circulatory System	31 (18.8)	5 (16.7)	21 (20.4)	5 (15.6)	
Diseases of Digestive System	18 (10.9)	2 (6.7)	9 (8.7)	7 (21.9)	
All other diagnoses	31 (18.8)	5 (16.7)	19 (18.4)	7 (21.9)	
SDU Admission Source					0.054
Higher intensity monitoring unit	72 (43.6)	19 (63.3)	41 (39.8)	12 (37.5)	
Direct Admission, or lower intensity monitoring unit	93 (56.4)	11 (36.7)	62 (60.2)	20 (62.5)	
Charlson-Deyo Comorbidity Index					0.008 ^{**}
0	74 (44.8)	18 (60.0)	36 (35.0)	20 (62.5)	
1	43 (26.1)	3 (10.0)	32 (31.1)	8 (25.0)	
> 2	48 (29.1)	9 (30.0)	35 (34.0)	4 (12.5)	

^{*} Oneway ANOVA p-value;

^{**} Chi square test p-value

Outcomes between clusters of patients who developed cardiorespiratory instability event (CRI₁) during the course of their stay on the step-down unit (SDU) length of stay (LOS) by cluster

Table 4

Variable	Total (N = 165)	Cluster 1 (n = 30)	Cluster 2 (n = 103)	Cluster 3 (n = 32)	p-value
Hospital LOS (days), mean, SD	13.5 ± 23	25.2 ± 47	10.6 ± 10.6	11.7 ± 10.3	0.006 *
SDU LOS (days), mean, SD	4.8 ± 4.7	5.2 ± 5.1	4.5 ± 4.2	5.5 ± 5.6	0.493
Transferred from SDU to a higher intensity monitoring unit	30 (18.2)	7 (23.3)	20 (19.4)	3 (9.4)	0.315
n (%)					

* Oneway ANOVA p-value