Heart & Lung 61 (2023) 107–113

Contents lists available at ScienceDirect

Heart & Lung

journal homepage: www.heartandlung.com

Unsupervised machine learning identifies symptoms of indigestion as a predictor of acute decompensation and adverse cardiac events in patients with heart failure presenting to the emergency department

Karina Kraevsky-Phillips, PhD(c), RN[a,b], Susan M. Sereika, PhD[a], Zeineb Bouzid, PhD(c)[a], Gavin Hickey, MD[b], Clifton W. Callaway, MD, PhD[a,b], Samir Saba, MD[a,b], Christian Martin-Gill, MD[a,b], Salah S. Al-Zaiti, PhD[a,*]

[a] University of Pittsburgh, Pittsburgh, PA, USA
[b] University of Pittsburgh Medical Center (UPMC), Pittsburgh, PA, USA

A R T I C L E I N F O

Article History:
Received 27 March 2023
Revised 12 May 2023
Accepted 21 May 2023
Available online 27 May 2023

Keywords:
Heart failure
Chest pain
Indigestion
Dyspnea
Symptom clusters
Machine learning

A B S T R A C T

Background: Patients with known heart failure (HF) present to emergency departments (ED) with a plethora of symptoms. Although symptom clusters have been suggested as prognostic features, accurately triaging HF patients is a longstanding challenge.
Objectives: We sought to use machine learning to identify subtle phenotypes of patient symptoms and evaluate their diagnostic and prognostic value among HF patients seeking emergency care.
Methods: This was a secondary analysis of a prospective cohort study of consecutive patients seen in the ED for chest pain or equivalent symptoms. Independent reviewers extracted clinical data from charts, including nine categories of subjective symptoms reported during initial evaluation. The diagnostic outcome was acute HF exacerbation and prognostic outcome was 30-day major adverse cardiac events (MACE). Outcomes were adjudicated by two independent reviewers. K-means clustering was used to derive latent patient symptom clusters, and their associations with outcomes were assessed using multivariate logistic regression.
Results: Sample included 438 patients (age 65±14 years; 45% female, 49% Black, 18% HF exacerbation, 32% MACE). K-means clustering identified three presentation phenotypes: patients with dyspnea only (Cluster A, 40%); patients with indigestion, with or without dyspnea (Cluster B, 23%); patients with neither dyspnea nor indigestion (Cluster C, 37%). Compared to Cluster C, indigestion was a significant predictor of acute HF exacerbation (OR=1.8, 95%CI=1.0–3.4) and 30-day MACE (OR=1.8, 95%CI=1.0–3.1), independent of age, sex, race, and other comorbidities.
Conclusion: Indigestion symptoms in patients with known HF signify excess risk of adverse events, suggesting that these patients should be triaged as high-risk during initial ED evaluation.

Introduction

More than six million people in the United States have heart failure (HF).[1] There are approximately one million incident cases diagnosed annually, and it is projected that the prevalence of HF will increase by 46% by 2030.[2] Despite progressive advances in guideline-directed medical therapies, survival rates have leveled off over time, with the current 5-year mortality rate remaining as high as 42.3%–52.6%.[3,4] In fact, HF is listed as a cause of death in approximately 13% of death certificates nationwide.[3] Given such a high morbidity burden, there are 1.4 million visits to the emergency department (ED) for HF annually, and nearly 20% of hospital admissions from the ED are due to HF.[5]

HF is a progressive, debilitating and highly symptomatic illness, but can result in varying presentation when individuals seek care. On the one hand, the symptoms can be persistent and non-specific (e.g., fatigue, loss of appetite), leading to symptom minimization, avoidance, denial, normalization, and improper attribution, especially in an aging population suffering from comorbid conditions.[6,7] Alternatively, symptoms can also be acute and troublesome (e.g., dyspnea, chest pain), but not specific for the underlying etiology (e.g.,

List of abbreviations: HF, heart failure; ED, emergency department; ECG, electrocardiogram; EHR, electronic health record; PCI, percutaneous coronary intervention; MACE, major acute cardiac events; AHA, American Heart Association; ACC, American College of Cardiology; ICD, implantable cardioverter defibrillator; ANOVA, analysis of variance; PCA, principal component analysis; ACS, acute coronary syndrome; IQR, interquartile range; OR, odds ratio; GI, gastrointestinal; NLP, natural language processing

* Corresponding author.
E-mail address: ssa33@pitt.edu (S.S. Al-Zaiti).

https://doi.org/10.1016/j.hrtlng.2023.05.012
0147-9563/

infarction, acute HF exacerbations, anxiety), leading to repeat ED visits and hospitalizations.[8,9] Opportunely, growing evidence suggests that HF symptoms tend to occur in clusters that have associations with event-free survival and other outcomes,[10-12] and can be used as a vital element in early clinician assessment and management of individuals with HF.[13]

While there are numerous standard ED risk assessment scores, they have shortcomings when it comes to the HF population. For instance, the HEART score is typically used for rapid adverse event risk assessment, diagnosis, and treatment decisions at time of patient presentation with complaints of chest pain. The score, which ranges from 0 to 10 (smallest to highest risk) is aggregated based on History, ECG, Age, Risk factors, Troponin and is assessed once at onset of encounter.[14] Similarly, the Emergency Severity Index (ESI)[15] is the standard for calculating patient illness severity based on chief complaint and patient presentation, factoring in estimated resource utilization. While both scores are routinely used, neither are tailored specifically to patients with HF. Since HF is a chronic illness with episodic decompensations that result in over one million ED visits annually, it is important to find methods of early identification of patients at high risk for poor outcomes, such as those experiencing acute decompensation, for effective triage and timely intervention.

With potential for predicting outcomes, HF symptom clusters may present an untapped resource at initial ED encounter. Clinical decisions in the ED are heavily based on initial evaluation. Additionally, unlike during a hospitalization, where symptoms change with progression of treatment, symptoms seen in the ED reflect fewer manifestations of underlying etiology and are not influenced by changes in presentation due to treatment. This plays a crucial role in early hospital management strategies and determinations about care pathways.[13] However, despite frequent ED visits for symptoms, there is a paucity of research exploring application of symptom clusters in ED setting. Thus, we sought to leverage recent advances in machine learning to study the diagnostic and prognostic value of symptom clusters in patients with HF seeking emergency care. We investigated whether unsupervised machine learning could identify clusters of symptoms associated with important outcomes in patients with HF presenting to an ED with chest pain. We posit that elucidating such clusters present at time of triage can help clinicians, especially ED triage nurses who are typically the first to encounter a patient and collect symptom data, to identify high priority patients and improve differentiation between patients reporting similar chief complaints.

Methods

Design, sample, and setting

This was a secondary analysis of a prospective observational cohort study. We describe the methods of the parent study in detail elsewhere[16-18] and the parent study is registered on clinicaltrials.gov (#NCT04237688). Briefly, the parent study recruited consecutive adult patients seeking care for chest pain or equivalent symptoms in the EDs of three UPMC affiliated tertiary care hospitals in the United States between 2013 and 2020 (n = 4132). We recruited all consecutive patients >18 years old arriving by ambulance with a high degree of clinical suspicion of a cardiac etiology requiring a prehospital acquisition of a 12-lead electrocardiogram (ECG). There were no specific exclusion criteria based on demographic or acuity of illness (e.g., cardiac arrest). This secondary analysis only included patients with a known history of HF as documented in electronic health records (EHR) at the time of ED visit (n = 438). The Institutional Review Board of the University of Pittsburgh approved a waiver of informed consent for minimal risk research, allowing us to enroll all eligible patients with no selection bias based on age, sex, race, or acuity.

Study variables

Independent reviewers manually extracted variables from EHR for the parent study. In addition to automatically populated data from prior encounters, EHRs are also updated every visit by the admitting clinician as part of routine care, and via interface with other databases, ensuring information accuracy. The extracted variables included baseline demographic and clinical characteristics, vital signs, results of diagnostic testing, and interventional and surgical treatments during hospitalization. We defined data elements in compliance with American Heart Association (AHA) and American College of Cardiology (ACC) definitions for assessing the clinical management and outcomes of patients with suspected acute coronary syndromes.[19] We limited potential independent variables for this analysis to include only symptom-related data as well as patient demographics, and past medical history − data that should be easily accessible at the time of triage in the ED.

Symptom data sources comprised clinical notes, including ED summary reports and nursing notes. The collected symptom data focused on general symptoms that are typically used in ED triage and would be important for patients with chest pain complaints.[20] We coded binary (yes/no) patient-reported symptoms into nine categories: (1) chest pain (chest discomfort, pressure, tightness, or other relevant patient descriptors); (2) dyspnea (dyspnea, breathlessness, or difficulty breathing); (3) indigestion (GI symptoms, fullness, bloating, nausea, vomiting, or upset stomach); (4) faintness (syncope, pre-syncope, dizziness, or lightheadedness); (5) heart rhythm complaints (palpitations, skipped heartbeat, pacemaker malfunction, or ICD shocks); (6) diaphoresis or sweating; (7) cough and other respiratory complaints (flu-like symptoms, or lower respiratory complaints); (8) fatigue (malaise, or general weakness); and (9) mental health complaints (anxiety, stress, panic attack, or substance use disorders [i.e., cocaine, alcohol intoxication]).

Two independent reviewers adjudicated the primary outcomes in the parent study based on a comprehensive manual chart review. The diagnostic outcome for this study was diagnosis of acute HF exacerbation as the suspected cause of symptoms that prompted patient to seek emergency care, defined as documented acute decompensation of preexisting HF; or worsening of end-stage/advanced HF associated with left ventricular systolic dysfunction.[21] The primary prognostic outcome was 30-day major adverse cardiac events (MACE), which is commonly used in cardiovascular research and with some variability incorporates important clinical endpoints such as revascularization and HF exacerbation.[22,23] For this study we defined it as a composite endpoint of cardiac arrest, ventricular tachycardia/fibrillation, new acute HF exacerbation event, cardiogenic shock, mechanical ventilation, or all-cause death occurring within 30 days of discharge from the indexed encounter, either from ED or hospitalization. On a randomly selected subset from the parent study (n = 165), the agreement between the two reviewers was high with a kappa coefficient of 0.91 (i.e., substantial to perfect agreement).

Statistical analysis

For all continuous variables we assessed normality, central tendency, standard deviation, and outliers. For nominal variables, all of which were measured on a binary scale, we examined frequencies and percentiles. All data were manually curated, alleviating issues with missingness. We characterized the sample using baseline demographic and clinical data available during initial ED evaluation, reporting mean ± standard deviation (SD) for continuous variables and frequency (%) for categorical variables. We compared groups using independent samples t-test or analysis of variance (ANOVA) for continuous variables, or chi-square test of independence for categorical variables. We used Spearman's rank-order correlation coefficient for studying bivariate associations between the binary-coded symptoms. We applied a Bonferroni correction for multiple comparisons and a heatmap to display the correlation matrix.

To identify latent groups and clusters in symptom data, we used principal component analysis (PCA) and k-means clustering. PCA compresses columns and k-means clustering compresses rows. For example, while PCA may conclude that "cough" accompanies "dyspnea," k-means clustering groups patients based on similarity in presentation (e.g., patient A and B are similar and belong to the same group). Nevertheless, both are unsupervised machine learning techniques that can reveal hidden patterns in the data without mapping input variables to an output class.

Our input data were a binary 0/1 matrix of symptom data in the shape of 438 × 9 (rows × columns). The PCA algorithm is a dimensionality reduction technique that searches for linear mathematical combinations (eigenvalues) that can reduce the original number of variables (i.e., nine symptoms) into a smaller subset of variables (i.e., principal components) while preserving the information in the matrix (i.e., the variability in the data).[24] For this technique, we also applied a varimax orthogonal rotation to transform and simplify the expression of the components. Extraction of the components was based on eigenvalues greater than 1. On the other hand, k-means clustering is an intuitive algorithm that maps individual patients in the dataset (i.e., 438 rows) to different clusters in the n-dimensional space in a way that the Euclidian distances between each point and the neighboring centroids of the other clusters are maximized. For this technique, we used the silhouette analysis as a diagnostic method to select the optimal number of clusters that maximizes the separation between patient clusters.[25] *A priori*, our sample size ($n = 438$) was adequate for both PCA[26] and k-means[27] clustering. We performed unsupervised learning approaches using Python v3.9 (Anon, 2020. Anaconda Software Distribution, Anaconda Inc.).

Finally, as we were interested in presence or absence of outcome, we used binary logistic regression to identify the univariate and multivariate predictors of the primary diagnostic and prognostic study outcomes. We used the resulting cluster assignment from k-means clustering as a categorical predictor in logistic regression models. Multivariate models included predictor variables significant at $p<0.10$ in univariate analysis to identify a parsimonious set of independent predictors. For the latter, we estimated and reported the adjusted odds ratio (OR) with 95% confidence interval. We conducted all statistical analyses using IBM® SPSS® Statistics (Version 28.0, Armonk, NY), and defined statistical significance as a *p*-value of 0.05 for two-sided hypothesis testing.

Results

The study sample included 438 patients (age 65±14 years, 45% female, 49% Black, 15% confirmed acute coronary syndrome [ACS]). Table 1 shows the demographic and clinical characteristics collected during the initial ED encounter as well as primary study outcomes. On average, each patient had a median of 3 comorbidities (25th – 75th percentile interquartile range [IQR] = 2–4), most frequently hypertension (94%), known history of coronary artery disease (CAD) (67%), dyslipidemia (61%), and current or prior history of smoking (60%). The average patient acuity assessed by the HEART score[14] was approximately 5 on a scale of 0–10 (most benign to most severe) with 80.4% triaged at intermediate (HEART score 4–6) to high risk (HEART score 7–10). Among these patients, 18% were adjudicated as having HF exacerbation, and subsequently, 32% of patients experienced 30-day MACE.

Fig. 1 shows the frequency and prevalence of patients with HF experiencing each of the nine symptoms of interest. Each patient reported a median of two symptoms (IQR 2–4), and a median duration from symptom onset to ED presentation of three hours (IQR = 1–12). Chest pain and other angina-like complaints were the most frequently reported symptoms in this cohort, followed by dyspnea, indigestion, and faintness. Mental health complaints, fatigue, and cough were the most infrequently reported. Fig. 2 shows the correlation matrix between the nine different symptoms. The heatmap shows that patients with heart rhythm complaints (e.g., palpitation, defibrillator shocks) are less likely to report chest pain, while diaphoresis and indigestion as well as dyspnea and cough tend to co-exist. However, most other bivariate correlations between symptoms were negligible.

We explored two conceptually different clustering methods to test their performance. We chose to start with PCA since symptom cluster research studies frequently use it for analysis. However, given the lack of correlations between symptom data (Fig. 2), the PCA algorithm yielded numerous models with Kaiser-Meyer-Olkin (KMO) statistic < 0.60, indicating poorly performing models.[28] Thus, we proceeded with k-means clustering using data from all nine symptoms as the input matrix. Fig. 3 shows the silhouette analysis for diagnosing 2-, 3-, 4-, and 5-cluster solutions for the k-means clustering algorithm. As shown in the figure, the average silhouette scores were highest for a solution based on three clusters, suggesting that grouping patients into these clusters would provide the best separation between decision boundaries while maintaining roughly equal

Table 1
Baseline sample characteristics and patient outcomes.

Characteristic/Outcome	All Patients ($N = 438$)
Age (years)	65 ± 14 (22–96)
Female Sex	195 (45%)
Black Race	213 (49%)
Comorbidities	
Hypertension	413 (94%)
known CAD	293 (67%)
Dyslipidemia	263 (61%)
Ever Smoked	261 (60%)
Diabetes mellitus	209 (48%)
COPD	165 (38%)
Stroke	74 (17%)
ACUITY ASSESSMENT	
HEART SCORE	4.7 ± 1.3
LOW RISK (0–3)	86 (20%)
INTERMEDIATE RISK (4–6)	318 (73%)
HIGH RISK (7–10)	27 (6%)
Outcome	
Acute HF exacerbation	80 (18%)
30-day MACE	138 (32%)

Values in the table are mean ± SD (range) or n (%).

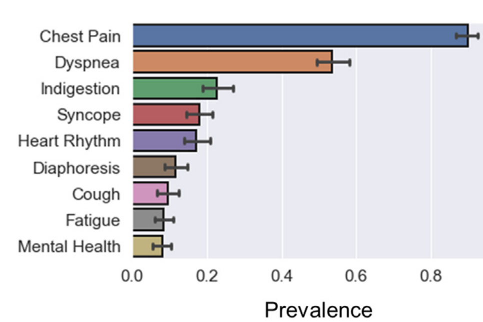

Fig. 1. Prevalence of ED Symptoms in Patients with Heart Failure
This is a bar graph showing the prevalence rate of presenting symptoms in heart failure patients evaluated in the emergency department. Symptoms are not mutually exclusive and are displayed in descending order of frequency. Chest pain broadly captures a wide range of subjective patient descriptors of discomfort, pressure, and tightness in the chest. GI includes feelings of indigestion, nausea, vomiting, and upset stomach. Heart rhythm complaints include palpitation, abnormal heart beats, and ICD malfunction or shocks. Mental health issues include anxiety, panic attack, and substance use disorders (e.g., cocaine abuse). Horizontal capped lines indicate 95%CI.

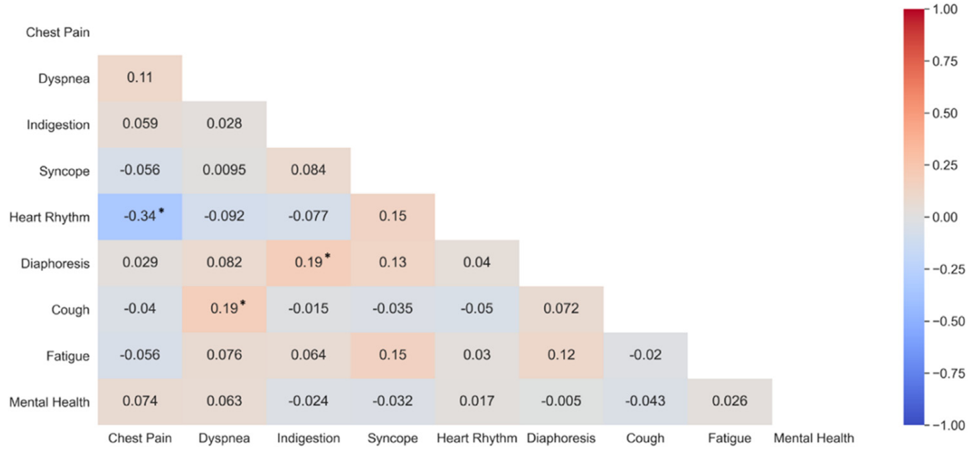

Fig. 2. Correlation Matrix of Patients' Symptoms

This is a correlation matrix of presenting symptoms with heatmap coding to indicate the strength of correlation. The warm/cool color coding in the legend key indicates positive vs. negative correlation, respectively. Values indicate bivariate correlation coefficients based on Spearman's r. Significance at $p < 0.05$ is denoted with asterisk after correction for multiple testing. Results suggest that patients with heart rhythm complaints (e.g., palpitation, defibrillator shocks) are less likely to report chest pain, while diaphoresis and indigestion as well as dyspnea and cough tend to co-exist.

patient distributions between the resulting clusters (i.e., Cluster A [n = 178]; Cluster B [n = 100]; and Cluster C [n = 160]). Fig. 4 describes the frequency and prevalence of symptoms in patients grouped in these clusters. Cluster A (dyspnea only cluster) exclusively consists of patients with dyspnea but no indigestion; Cluster B (indigestion cluster) exclusively consists of patients with indigestion, with or without dyspnea; and Cluster C (other symptoms cluster) consists of patients with neither dyspnea nor indigestion.

Table 2 reports the description and comparison of the baseline demographic and clinical characteristics among the three patient clusters. There were no differences among clusters in terms of age, sex, and most comorbidities. Cluster C (other symptoms cluster) was more likely to include patients who are Black, and less likely to include patients who ever smoked or those with COPD. Despite a similar acuity profile at ED presentation using HEART score (F [2, 435] = 0.10, p = 0.905), patients in Cluster A (dyspnea cluster) and Cluster B (indigestion cluster) were more likely to experience 30-day MACE, and patients in Cluster B (indigestion cluster) were more likely to be diagnosed with HF exacerbation. In multivariate binary logistic regression analysis controlling for selected baseline demographic and clinical characteristics and patient acuity (Table 3), compared to patients in Cluster C (other symptoms cluster), patients in Cluster A (dyspnea cluster) did not have higher odds of experiencing HF exacerbation but were at excess risk of 30-day MACE (OR = 1.6, 95% CI 1.0–2.6). Compared to patients in Cluster C (other symptoms cluster), patients in Cluster B (indigestion cluster) were at excess risk of both experiencing HF exacerbation (OR = 1.8, 95% CI 1.0–3.4) and 30-day MACE (OR = 1.8, 95% CI 1.0–3.1).

Discussion

The specific aims of this study were to examine the diagnostic and prognostic value of symptom clusters in patients with HF seeking emergency care for chest pain or its equivalent. Our results showed that the two most important symptoms to define symptom clusters were dyspnea and indigestion. The absence of both symptoms, which is observed in 37% of patients, was a strong protective factor that can indicate no acute decompensation and lower odds of 30-day adverse cardiac events. However, the presence of dyspnea or indigestion had important diagnostic and prognostic value. Most notably, the presence of symptoms of indigestion, with or without dyspnea, was a strong predictor of both ongoing acute decompensation and excess risk of 30-day adverse cardiac events. These symptom clusters were independent of initial acuity assessment in the ED and other baseline comorbidities, which can have important clinical implications during initial ED triage by nurses and physicians alike. Additionally, it was intriguing that patients with heart rhythm complaints were less likely to have concurrent chest pain (Fig. 2). Perhaps this was the result of the methodological design of the parent study that recruited patients with complaints of chest pain or equivalent symptoms. These equivalent symptoms would include other symptoms that prompt the acquisition of an ECG, such as palpitations, regardless of presence of chest pain.

Mechanisms of symptoms occurring concurrently with chest pain have been explored previously. Chest pain in general has been linked to gastrointestinal (GI) symptoms such as nausea and indigestion via mechanisms such as coronary spasm,[29] pain receptor innervation[30] and endothelial dysfunction.[31] Dyspnea with chest pain, on the other hand, has been posited to result from worsening diastolic dysfunction, leading to increased pulmonary congestion.[32] Notably, our findings corroborate with other studies that link dyspnea and GI symptoms with deterioration and poor outcomes in patients with HF. Dyspnea, a hallmark symptom of HF, has been shown to have varying physiological mechanisms ranging from increased ergoreflex sensitivity[33] to pulmonary congestion due to left-ventricular dysfunction,[34] making it difficult to attribute directly to any one physiological change in the body. Indigestion or GI distress in the HF population are a sign of poor intestinal blood flow,[35] and most importantly low output HF and congestion.[36] Thus, GI symptoms are a potential indicator of existing hypoperfusion and increased congestion, carry significant clinical implications, including increased risk for poor outcomes,[37] and require more thorough investigation. However, regardless of the exact mechanism responsible for the perceived symptoms, occurrence of both indigestion and dyspnea in symptom clusters have been correlated with poor outcomes both in our study and elsewhere. For example, Jurgens, Lee, and Riegel[10] reported that the dyspnea-related cluster and cluster that included many GI-related symptoms were associated with increased risk of HF-related adverse events at one year.

The components that were extracted by PCA did not perform strongly. While there were correlations between some symptoms, they were not strong enough to produce meaningful "clustering" results. One potential explanation for this finding is the binary nature of the symptom data which has limited variability and dimensionality. Other HF symptom cluster studies used a Likert scale questionnaire for collecting symptom data, which allowed for expanded

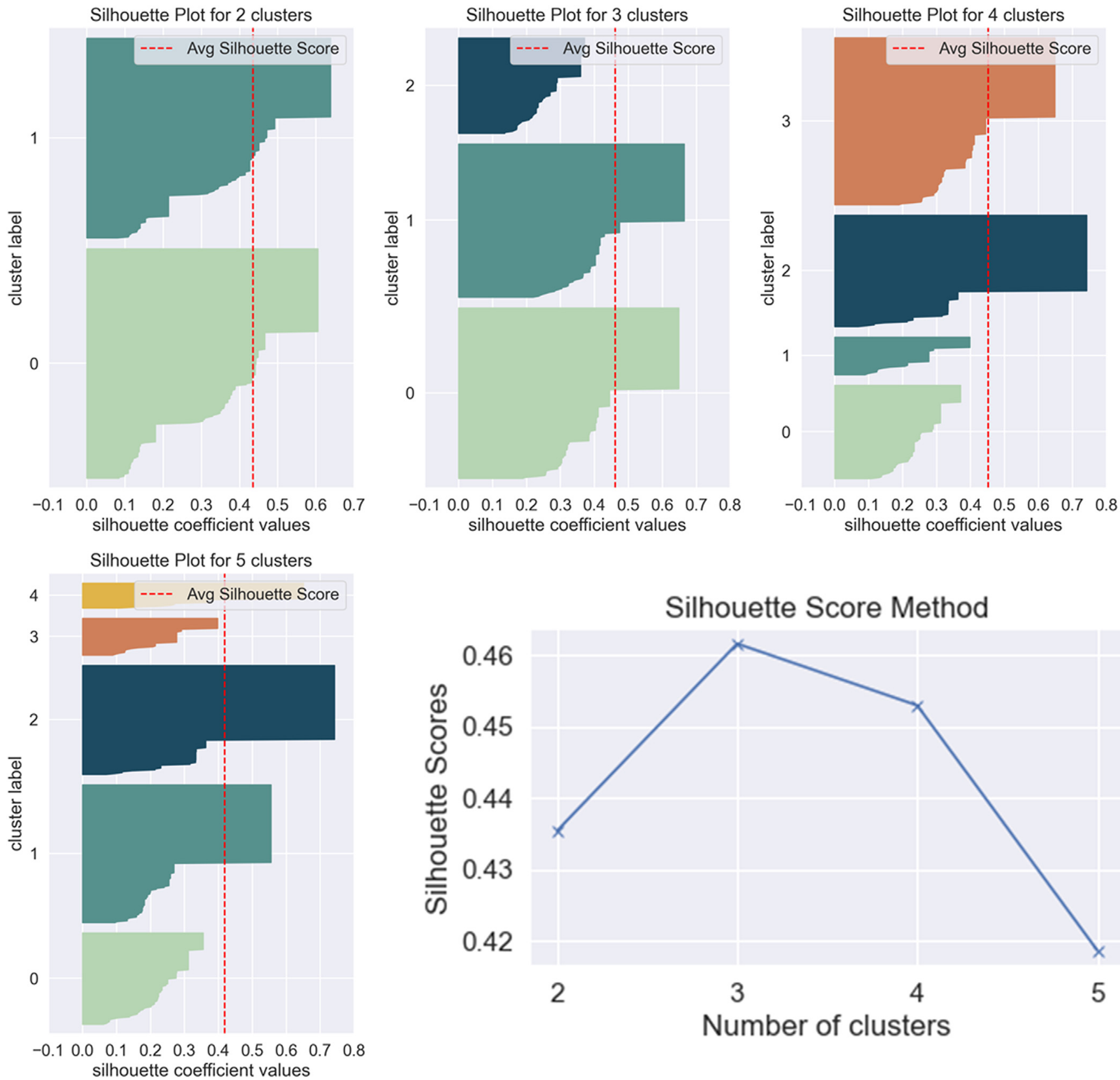

Fig. 3. K-means clustering diagnostics using silhouette analysis

These silhouette plots display the separation between each point in each cluster to the neighboring clusters when a k-means solution is optimized for 2, 3, 4, or 5 clusters. The x-axis in each plot shows the silhouette coefficients based on normalized Euclidian distances where +1 indicates large separation from neighboring clusters and 0 indicates a point is very close to the decision boundary between two neighboring clusters. The average silhouette scores are denoted with dashed lines in each plot and are summarized in the line plot in the right lower corner. This silhouette analysis suggests that using 3 clusters provides the best separation while maintaining roughly equal patient distributions (thickness of plots) between the resulting clusters.

information about symptom severity,[10,11] and resulted in data more suitable for PCA. The present study was based on symptoms extracted from tabular data for clinical parameters whose severity level is not typically documented in EHRs like that of pain. One potential future solution for this problem is application of natural language processing (NLP) to mine symptom severity information about symptoms since narrative notes would be more likely to contain descriptive data.

This study has numerous clinical implications, most important of which is treatment and disposition decisions. Evidence suggesting the importance of early treatment exists, which includes timely assessment by triage nurses.[38] Nurses are trained to recognize important signs and symptoms when patients present to the ED and will often be the first clinician encountered during an ED visit.[13] This study showed potential value in easily discernable patient characteristics, such as symptoms, to make treatment decisions. Patients with HF typically have multiple coexisting comorbidities that can lead to similar symptoms. Thus, many of the symptoms reported as chief complaints by patients can overlap across multiple possible disease processes, making treatment and discharge decisions challenging. Furthermore, studies have shown that patients presenting with dyspnea may be challenging to diagnose, with only approximately 80% accuracy in identifying acute heart failure,[39] which may result in inappropriate treatment and higher mortality.[40] Findings of this

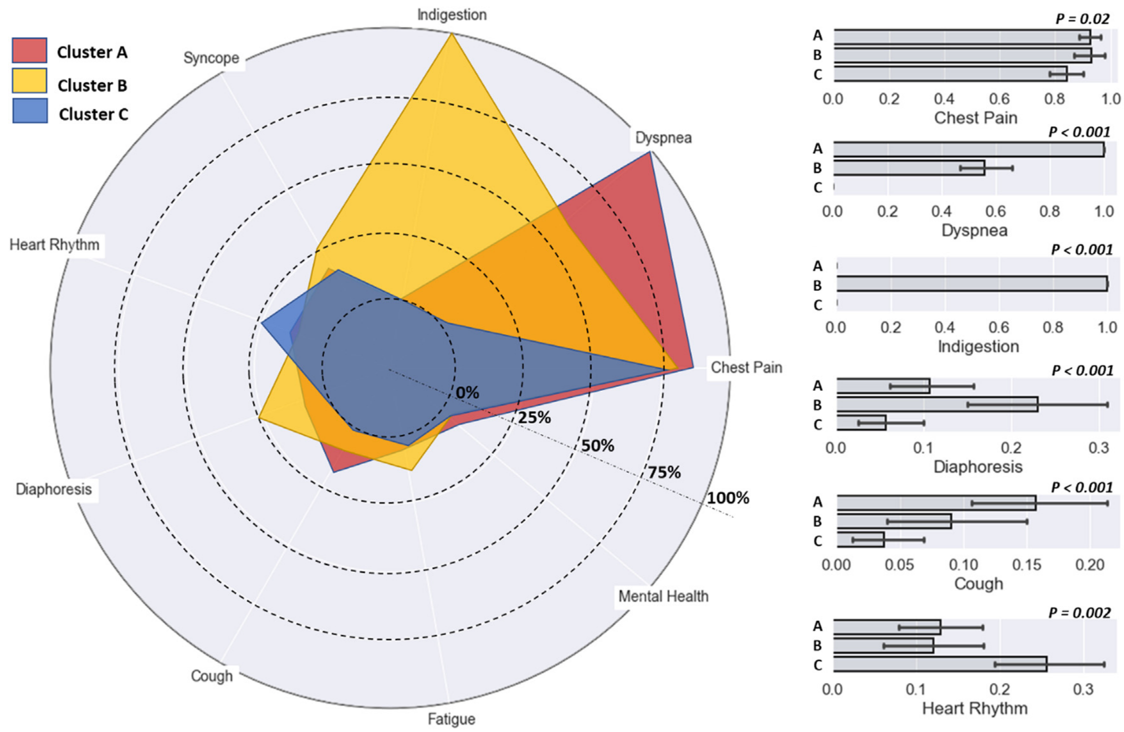

Fig. 4. Relations of Patient Symptoms to K-means Clusters

The left panel shows the prevalence of symptoms per each k-means cluster visualized with a radar chart. Dashed inner circles with corresponding percentages denote the reference axis for prevalence rate. The right panel shows individual bar charts with statistically significant associations between a given symptom and k-means clusters. These two panels indicate that Cluster A consists of patients with chest pain AND dyspnea but no GI symptoms; Cluster B consists of patients with chest pain AND indigestion with or without dyspnea or diaphoresis; and Cluster C consists of patients with chest pain but no dyspnea or indigestion.

Table 2
Comparison of baseline characteristics and outcomes between k-means clusters.

Characteristic	Cluster A (n = 178)	cluster b (n = 100)	cluster c (n = 160)
Age (years)	65 ± 15	65 ± 14	67 ± 15
Female Sex	80 (45%)	51 (51%)	64 (40%)
Black Race	**83 (46%)**	**42 (42%)**	96 (60%)
Comorbidities			
Hypertension	167 (94%)	93 (93%)	153 (96%)
Known CAD	119 (67%)	65 (65%)	109 (68%)
Dyslipidemia	109 (61%)	65 (65%)	93 (58%)
Ever Smoked	**117 (66%)**	**61 (61%)**	83 (52%)
Diabetes mellitus	93 (52%)	48 (48%)	68 (43%)
COPD	**74 (42%)**	**45 (45%)**	46 (29%)
Stroke	33 (19%)	13 (13%)	28 (18%)
ACUITY ASSESSMENT			
HEART SCORE	4.7 ± 1.2	4.7 ± 1.3	4.7 ± 1.5
Outcome			
Acute HF exacerbation	32 (18%)	**25 (25%)**	23 (14%)
MACE	**61 (34%)**	**38 (38%)**	39 (24%)

Cluster A, dyspnea without indigestion; Cluster B, indigestion, with or without dyspnea; Cluster C, neither dyspnea nor indigestion. Bold indicates significance against "Cluster C" as a reference.

Table 3
Univariate and multivariate predictors of primary outcomes.

	Acute hf exacerbation[†] Multivariate OR (95% CI)	30-day mace[‡] Multivariate OR (95% CI)
Age (Years)	1.0 (0.99–1.01)	1.0 (0.99–1.02)
Male Sex	1.5 (0.9–2.5)	1.3 (0.8–2.1)
ever smoked	**1.7 (1.0–2.9)**	**1.5 (1.0–2.4)**
COPD	**1.7 (1.0–2.9)**	1.0 (0.6–1.6)
HEART score	**1.2 (1.0–1.5)**	**1.4 (1.2–1.7)**
k-means clusters		
cluster a	1.2 (0.6–2.1)	**1.6 (1.0–2.7)**
cluster b	**1.8 (1.0–3.5)**	**2.0 (1.1–3.5)**
cluster c	Ref	Ref

Bold indicates statistical significance at $p < 0.05$.
[†] Hosmer-and-Lemeshow chi-square (8) = 8.857, p = 0.354, R square = 0.068.
[‡] Hosmer-and-Lemeshow chi-square (8) = 5.378, p = 0.716, R square = 0.097.

study highlighted a low-risk group (~40%) with neither dyspnea nor indigestion. This subset of patients can be triaged as lower acuity where appropriate, treated, and discharged with lower concern for short-term adverse events. This study also identified a high-risk group (~25%) with symptoms of indigestion, with or without dyspnea, which may be a surrogate indicator for hypoperfusion and congestion. These patients could be prioritized during triage (to be seen first) and can receive more expedited testing (e.g., echocardiogram) in addition to the standard chest pain work up. This could potentially lead to faster treatment and better patient outcomes.

Our study had several limitations. First, while our sample had some favorable characteristics, the study was conducted in a small geographic area, thus limiting generalizability. Second, the binary nature of symptom data extracted from EHRs limited statistical analysis and made cluster detection more challenging. Finally, because of the setting of the parent study, likely only the most prevalent and troublesome symptoms (such as dyspnea) were reported, also limiting dimensionality of reported symptoms.

Conclusion

This study demonstrated feasibility of detecting symptom clusters in the ED setting. The derived clusters were predictive of clinically important outcomes. We demonstrated that the presence of indigestion in patients with HF presenting with chest pain was independently associated with acute decompensation and 30-day adverse cardiac events. Indigestion is an easy to assess phenotypic presentation, which has been previously linked with unique mechanisms of disease. Symptoms are readily available during initial patient triage in emergency departments and can be useful to clinicians when making treatment and disposition decisions. This is, to our knowledge, the first study examining symptom clusters in the ED. Further studies are needed to replicate the findings.

Funding

R01HL137761 (Al-Zaiti) and T32NR008857, ARCS Scholar (Kraevsky-Phillips)

Trial registration

ClinicalTrials.gov # NCT04237688

Declaration of Competing Interest

The authors have no conflicts of interest to disclose.

References

1. Savarese G, Lund LH. Global public health burden of heart failure. *Cardiac Failure Review*. 2017;03:7. https://doi.org/10.15420/cfr.2016:25:2.
2. Ambrosy AP, Fonarow GC, Butler J, et al. The global health and economic burden of hospitalizations for heart failure: lessons learned from hospitalized heart failure registries. *J Am Coll Cardiol*. 2014;63:1123–1133. https://doi.org/10.1016/j.jacc.2013.11.053.
3. Virani SS, Alonso A, Aparicio HJ, et al. Heart disease and stroke statistics—2021 update. *Circulation*. 2021:143. https://doi.org/10.1161/cir.0000000000000950.
4. Benjamin EJ, Virani SS, Callaway CW, et al. Heart disease and stroke statistics-2018 update: a report from the American Heart Association. *Circulation*. 2018;137:e67–e492. https://doi.org/10.1161/cir.0000000000000558.
5. Trends in the utilization of emergency department services, 2009-2018. 2021.
6. Bruce Wirta S, Balas B, Proenca CC, et al. Perceptions of heart failure symptoms, disease severity, treatment decision-making, and side effects by patients and cardiologists: a multinational survey in a cardiology setting. *Ther Clin Risk Manag*. 2018;14:2265–2272. https://doi.org/10.2147/tcrm.S183200.
7. Schumacher C, Hussey L, Hall V. Heart failure self-management and normalizing symptoms: an exploration of decision making in the community. *Heart Lung*. 2018;47:297–303. https://doi.org/10.1016/j.hrtlng.2018.03.013.
8. Alpert CM, Smith MA, Hummel SL, Hummel EK. Symptom burden in heart failure: assessment, impact on outcomes, and management. *Heart Fail Rev*. 2017;22:25–39. https://doi.org/10.1007/s10741-016-9581-4.
9. Riegel B, Jaarsma T, Lee CS, Strömberg A. Integrating symptoms into the middle-range theory of self-care of chronic illness. *Adv Nurs Sci*. 2019;42:206–215. https://doi.org/10.1097/ans.0000000000000237.
10. Jurgens CY, Lee CS, Riegel B. Psychometric analysis of the heart failure somatic perception scale as a measure of patient symptom perception. *J Cardiovasc Nurs*. 2017;32:140–147. https://doi.org/10.1097/JCN.0000000000000320.
11. Lee KS, Song EK, Lennie TA, et al. Symptom clusters in men and women with heart failure and their impact on cardiac event-free survival. *J Cardiovasc Nurs*. 2010;25:263–272. https://doi.org/10.1097/JCN.0b013e3181cfbb88.
12. Son YJ, Won MH. Symptom clusters and their impacts on hospital readmission in patients with heart failure: a cross-sectional study. *Res Theory Nurs Pract*. 2018;32:311–327. https://doi.org/10.1891/1541-6577.32.3.311.
13. Mebazaa A, Yilmaz MB, Levy P, et al. Recommendations on pre-hospital & early hospital management of acute heart failure: a consensus paper from the Heart Failure Association of the European Society of Cardiology, the European Society of Emergency Medicine and the Society of Academic Emer. *Eur J Heart Fail*. 2015;17:544–558. https://doi.org/10.1002/ejhf.289.
14. Six AJ, Backus BE, Kelder JC. Chest pain in the emergency room: value of the HEART score. *Netherland Heart J*. 2008;16:191–196. https://doi.org/10.1007/bf03086144.
15. Shelton R. The emergency severity index 5-level triage system. *Dimens Crit Care Nurs*. 2009;28:9–12. https://doi.org/10.1097/01.Dcc.0000325106.28851.89.
16. Al-Zaiti SS, Martin-Gill C, Sejdić E, Alrawashdeh M, Callaway C. Rationale, development, and implementation of the electrocardiographic methods for the prehospital identification of Non-ST elevation myocardial infarction events (EMPIRE). *J Electrocardiol*. 2015;48:921–926. https://doi.org/10.1016/j.jelectrocard.2015.08.014.
17. Al-Zaiti S, Besomi L, Bouzid Z, et al. Machine learning-based prediction of acute coronary syndrome using only the pre-hospital 12-lead electrocardiogram. *Nat Commun*. 2020;11:3966. https://doi.org/10.1038/s41467-020-17804-2.
18. Bouzid Z, Faramand Z, Gregg RE, et al. In search of an optimal subset of ECG features to augment the diagnosis of acute coronary syndrome at the emergency department. *J Am Heart Assoc*. 2021;10: e017871. https://doi.org/10.1161/jaha.120.017871.
19. Cannon CP, Brindis RG, Chaitman BR, et al. 2013 ACCF/AHA key data elements and definitions for measuring the clinical management and outcomes of patients with acute coronary syndromes and coronary artery disease. *Circulation*. 2013;127:1052–1089. https://doi.org/10.1161/cir.0b013e3182831a11.
20. Zegre-Hemsey JK, Crandell J, Rosamond WD, et al. Stable versus dynamic cardiac symptom characteristics are associated with adverse outcomes among individuals transported by ambulance for suspected acute coronary syndrome. *Circulation*. 2022:146.
21. Nieminen MS, Harjola VP. Definition and epidemiology of acute heart failure syndromes. *Am J Cardiol*. 2005;96. https://doi.org/10.1016/j.amjcard.2005.07.015. 5 g-10g.
22. Al-Zaiti SS, Faramand Z, Alrawashdeh MO, Sereika SM, Martin-Gill C, Callaway C. Comparison of clinical risk scores for triaging high-risk chest pain patients at the emergency department. *Am J Emerg Med*. 2019;37:461–467. https://doi.org/10.1016/j.ajem.2018.06.020.
23. Bosco E, Hsueh L, Mcconeghy KW, Gravenstein S, Saade E. Major adverse cardiovascular event definitions used in observational analysis of administrative databases: a systematic review. *BMC Med Res Methodol*. 2021;21. https://doi.org/10.1186/s12874-021-01440-5.
24. Ringnér M. What is principal component analysis? *Nat Biotechnol*. 2008;26:303–304. https://doi.org/10.1038/nbt0308-303.
25. Wang F, Franco-Penya H-H, Kelleher JD, Pugh J, Ross R. An analysis of the application of simplified silhouette to the evaluation of k-means clustering validity. *Machine Learning and Data Mining in Pattern Recognition*. Springer International Publishing; 2017:291–305.
26. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. eds. Boston: Pearson; 2019.
27. Dalmaijer ES, Nord CL, Astle DE. Statistical power for cluster analysis. *BMC Bioinformat*. 2022:23. https://doi.org/10.1186/s12859-022-04675-1.
28. Kaiser HF. An index of factorial simplicity. *Psychometrika*. 1974;39:31–36. https://doi.org/10.1007/BF02291575.
29. Chauhan A, Petch MC, Schofield PM. Effect of oesophageal acid instillation on coronary blood flow. *Lancet*. 1993;341:1309–1310. https://doi.org/10.1016/0140-6736(93)90817-z.
30. Castell DO. Chest pain of undetermined origin: overview of pathophysiology. *Am J Med*. 1992;92. https://doi.org/10.1016/0002-9343(92)80049-6. 2s-4s.
31. Oparin A, Vnukova A. The role of endothelial dysfunction in the mechanism of gastroesophageal reflux disease development in patients with ischemic heart disease. *Acta Clin Croat*. 2017;56:635–639. https://doi.org/10.20471/acc.2017.56.04.08.
32. Marcus G, Shimony S, Stein GY, et al. Incidence, predictors and prognostic implications of dyspnea at admission among acute coronary syndrome patients without heart failure. *Int J Cardiol*. 2020;301:29–33. https://doi.org/10.1016/j.ijcard.2019.11.093.
33. Aimo A, Saccaro LF, Borrelli C, et al. The ergoreflex: how the skeletal muscle modulates ventilation and cardiovascular function in health and disease. *Eur. J. Heart Fail.*. 2021;23:1458–1467. https://doi.org/10.1002/ejhf.2298.
34. Kupper N, Bonhof C, Westerhuis B, Widdershoven J, Denollet J. Determinants of dyspnea in chronic heart failure. *J Card Fail*. 2016;22:201–209. https://doi.org/10.1016/j.cardfail.2015.09.016.
35. Valentova M, von Haehling S, Bauditz J, et al. Intestinal congestion and right ventricular dysfunction: a link with appetite loss, inflammation, and cachexia in chronic heart failure. *Eur Heart J*. 2016;37:1684–1691. https://doi.org/10.1093/eurheartj/ehw008.
36. Sundaram V, Fang JC. Gastrointestinal and liver issues in heart failure. *Circulation*. 2016;133:1696–1703. https://doi.org/10.1161/circulationaha.115.020894.
37. Javaloyes P, Miró Ò, Gil V, et al. Clinical phenotypes of acute heart failure based on signs and symptoms of perfusion and congestion at emergency department presentation and their relationship with patient management and outcomes. *Eur J Heart Fail*. 2019;21:1353–1365. https://doi.org/10.1002/ejhf.1502.
38. Milbrath G, Snyder A. The emergency nurses association: 50 years of advocacy and advancement. *J Emerg Nurs*. 2021;47:224–238. https://doi.org/10.1016/j.jen.2020.11.007. e221.
39. Steinhart BD, Levy P, Vandenberghe H, et al. A randomized control trial using a validated prediction model for diagnosing acute heart failure in undifferentiated dyspneic emergency department patients-results of the GASP4Ar study. *J Card Fail*. 2017;23:145–152. https://doi.org/10.1016/j.cardfail.2016.08.007.
40. Ray P, Birolleau S, Lefort Y, et al. Acute respiratory failure in the elderly: etiology, emergency diagnosis and prognosis. *Criti Care*. 2006;10:R82. https://doi.org/10.1186/cc4926.