Published in final edited form as:
*J Am Coll Cardiol*. 2014 October 28; 64(17): 1765–1774. doi:10.1016/j.jacc.2014.07.979.

# Clinical Implications of Chronic Heart Failure Phenotypes Defined by Cluster Analysis

**Tariq Ahmad, MD, MPH**[*,†], **Michael J. Pencina, PhD**[†], **Phillip J. Schulte, PhD**[†], **O'Brien Emily, PhD**[†], **David J. Whellan, MD**[‡], **Ileana L. Piña, MD, MPH**[§], **Dalane W. Kitzman, MD**[‖], **Kerry L. Lee, PhD**[†], **Christopher M. O'Connor, MD**[*,†], and **G. Michael Felker, MD, MHS**[*,†]

[*]Division of Cardiology, Duke University Medical Center

[†]Duke Clinical Research Institute, Durham, NC

[‡]Thomas Jefferson University, Philadelphia, PA

[§]Montefiore Medical Center, Bronx, NY

[‖]Wake Forest School of Medicine, Winston-Salem, NC

## Abstract

**BACKGROUND—**Classification of chronic heart failure (HF) is based on criteria that may not adequately capture disease heterogeneity. Improved phenotyping may help inform research and therapeutic strategies.

**OBJECTIVE—**This study used cluster analysis to explore clinical phenotypes in chronic HF patients.

**METHODS—**A cluster analysis was performed on 45 baseline clinical variables from 1,619 participants in HF-ACTION (Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training), evaluating exercise training versus usual care in chronic systolic HF. Association between identified clusters and clinical outcomes was performed using Cox proportional hazards modeling. Differential associations between clinical outcomes and exercise testing were examined using interaction testing.

**RESULTS—**Ranging in size from 248 to 773, four clusters were identified whose patients varied considerably along measures of age, sex, race, symptoms, comorbidities, HF etiology, socioeconomic status, quality of life, cardiopulmonary exercise testing parameters, and biomarker levels. Differential associations were observed for hospitalization and mortality risks between and within clusters. To illustrate, compared with cluster 1, risk of all-cause mortality/all-cause hospitalization ranged from 0.65 (0.54 to 0.78) for cluster 4 to 1.02 (0.87 to 1.19) for cluster 3.



**Address for correspondence**: G. Michael Felker, MD, MHS, Associate Professor of Medicine, Duke Clinical Research Institute, DUMC Box 3850, Durham, NC 27710, Phone: 919-668-8919, Fax: 919-668-7063, michael.felker@duke.edu.
**Disclosures:**None of the other authors report any conflicts.

However, for all-cause mortality, cluster 3 had disproportionately lower risk 0.61 (0.44 to 0.86). Evidence suggested differential effects of exercise treatment on changes in peak $VO_2$ and clinical outcomes between clusters (p for interaction <0.04).

**CONCLUSIONS—**Cluster analysis of clinical variables identified 4 distinct phenotypes of chronic HF. Our findings underscore the high degree of disease heterogeneity that exists within chronic HF patients and a need for improved phenotyping.

## Keywords

mortality; prognosis; rehospitalization; socioeconomic

Chronic heart failure (HF) is a syndrome rather than a specific disease, with several distinct subtypes that may respond uniquely to therapeutic interventions (1). However, despite advances in our understanding of HF pathogenesis, its classification continues to rely on imprecise measures that may lead to overlapping diagnostic labels and misclassification (2,3). For example, chronic HF is still clinically defined along subjective measures of functional status (New York Heart Association [NYHA] class), arbitrary left ventricular ejection fraction (LVEF) cut points (HF with preserved versus reduced EF), or stages (A to D), despite increasing recognition that these constructs provide inadequate phenotyping of the syndrome(4–6).

Inadequately classifying patients within a disease state like heart failure may produce several potentially important consequences. Since therapeutic interventions are frequently based on targeting certain patient subgroups, inadequate classification may lead to ineffective or inappropriate treatments. The shortcomings in contemporary HF classification have been posited as a possible explanation for why we have seen such little progress in developing new treatments for this disorder(7,8). Improving the ‘taxonomy’ of clinical classification may therefore offer important clinical benefits. Whereas molecular phenotyping might theoretically provide a more rational disease description, an essential first step is to identify disease sub-types based on key clinical variables, such that downstream biological measurements can be appropriately anchored in patient level data. Indeed, the National Research Council has released a report that calls for a new taxonomy of disease based on both clinical and molecular measures that will provide a more accurate classification of disease, with the ultimate goal of enhancing diagnosis and treatment (9).

A widely used exploratory and hypothesis-generating approach in biological studies, clustering has played important roles in identifying subtypes in complex diseases. This approach has been extensively used in analyzing molecular data across disease states, but seldom employed to examine clinical variables; however, several reports suggest that it can lead to improved characterization of disease phenotype (10,11). Accordingly, we applied cluster analysis to examine the presence of clinically important patient subgroups within a well-characterized cohort of chronic HF patients randomized to exercise training versus usual care. We also examined patterns of adverse clinical outcomes among derived patient clusters, as well as interaction with randomized treatment assignment.

# METHODS

## STUDY POPULATION

Details of the design, rationale, and primary results of HFACTION (Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training) have been published elsewhere (12,13). Briefly, HF-ACTION (clinicaltrials.gov: NCT00047437) was a randomized clinical trial evaluating the effect of exercise training versus usual care on long-term morbidity and mortality in 2,331 patients with chronic HF due to LV systolic dysfunction (NYHA class II to IV, LVEF ≤35%). Patients were randomized to either usual HF care or a structured, group-based, supervised exercise program. All patients, regardless of treatment group, received detailed self-management educational materials that included information on medications, fluid management, symptom exacerbation, sodium intake, and amount of activity recommended by ACC/AHA guidelines (14). Patients were followed for a median of 2.6 years.

## ASSESSMENT OF CLINICAL VARIABLES AND BIOMARKERS

At the baseline clinic visit prior to randomization, demographics, socioeconomic status, past medical history, current medications, a physical examination, and the most recent laboratory tests were obtained. Participants reported race and ethnicity at the time of study enrollment using categories defined by the National Institutes of Health. All patients underwent baseline and 3- month cardiopulmonary exercise testing (CPET), during which key exercise parameters were ascertained. Additionally, a standard 6-minute walk test (6MWD) was performed on each patient during the baseline visit. Transthoracic echocardiography (TTE) was performed at baseline and key measures acquired by the core laboratory included LVEF and mitral regurgitation assessment. Health status measures were ascertained using several validated psychometric instruments at baseline to measure health-related quality-of–life (QOL), pain, depression, and social support, including the Kansas City Cardiomyopathy Questionnaire (KCCQ), and the Multidimensional Scale of Perceived Social Support (15). Baseline biomarker levels of N terminal pro-B-type natriuretic peptide (NT-proBNP), ST2, and galectin-3 were evaluated in a subset of patients who agreed to participate in the biomarker substudy, using previously described methodologies (16,17).

## CLINICAL ENDPOINTS

The primary endpoint of HF-ACTION was a composite of allcause mortality and all-cause hospitalization over a median follow-up of 2.6 years. Additional endpoints of interest included change from baseline in peak oxygen consumption per unit of time (peak $VO_2$) at 3 months, all-cause mortality, a composite endpoint of cardiovascular (CV) mortality or CV hospitalization, and the composite endpoint of CV mortality or HF hospitalization. An independent clinical events committee adjudicated all deaths and all first hospitalizations.

## STATISTICAL ANALYSIS

Cluster analysis defines the distances between subjects based on the combined values of their measured characteristics. Using a matrix of distance measurements, cluster analysis finds groups of subjects more similar to each other than to those in other groups. It can be

used to describe disease phenotypes without the need for historical or arbitrary a priori assumptions about classification.

Details of the statistical analysis performed are included in the Supplementary Data section. Briefly, we selected 45 candidate variables measured at baseline that represented key characteristics of patients with HF, including demographics, medical history, laboratory values, QOL scores, and exercise capabilities (Supplementary Table 1). As is necessary for cluster analysis, patients with missing data for any variables were excluded, resulting in an analytic population of 1,619/2,331 (70% of the baseline study population). We performed a cluster analysis on these variables and obtained four distinct clusters of chronic HF patients. The association between cluster membership and clinical outcomes was assessed using Cox proportional hazards regression. We assessed proportional hazards assumptions graphically by evaluating the standardized score process and the supremum test and found no violations (18). Using interaction terms in a Cox regression model, we also assessed whether cluster membership was associated with a differential response to randomized exercise therapy for each outcome.

All analyses were performed with SAS 9.2 (SAS Institute Incorporated, Cary, NC) and R 2.15.3 (R Development Core Team, Vienna, Austria). A p value $\leq 0.05$ was considered statistically significant for all analyses. The authors had full access to and take full responsibility for data integrity.

## RESULTS

Complete baseline data for the pre-specified 45 clinical variables of interest were available for 1,619 of the 2,331 patients who participated in the HF-ACTION trial; these patients were included in the study. The cluster analysis identified 4 patient clusters; clinical variables of these clusters are shown in Table 1and socioeconomic variables in Table 2. Table 3 contains objective measures of HF according to patient cluster. Baseline characteristics of the overall population and the subgroup used for the analysis were broadly similar, and are shown in Supplementary Table 2. Key characteristics of each patient cluster were as follows:

### Cluster 1 (n=773)

This was the largest cluster with >2 times more patients than the other clusters. Patients tended to be older Caucasian males (>60 years) with a history of tobacco use, high rates of ischemic cardiomyopathy (68%), and advanced NYHA functional class (39% with class III or IV). Despite having the second highest rates of coronary artery bypass graft (CABG) surgery and percutaneous coronary intervention (PCI), they had the second lowest rates of angina symptoms (11.3%) with only 1% having Canadian Cardiovascular Society (CCS) Angina class 2 to 4. They had the highest rates of common co-morbidities such as atrial fibrillation (AF), renal insufficiency, and chronic obstructive pulmonary disease (COPD), as well as implantable cardioverter-defibrillators (ICDs) and coronary resynchronization therapy (CRT). Cluster 1 patients were most likely to be married, least likely to be divorced, had the second highest rates of college graduation and income, and were most likely to either be employed or retired (63%). They had objective evidence of the most advanced disease: lowest median peak $VO_2$ levels (13.5 ml/kg/min), highest ventilation versus

$CO_2$ production (VE-$VCO_2$) slope (34), and lowest 6MWD (351 meters), but they had the second lowest rates of prior HF hospitalization and the second highest QOL scores. They also had the highest median levels of all 3 HF biomarkers studied: NT-proBNP (1, 079 pg/ml), galectin-3 (15.4 ng/ml), and ST2 (26.2 ng/ml).

### Cluster 2 (n =287)

These patients were, on average, the youngest (median age = 49); most likely to be African Americans (69%), and had the second highest percentage of females (38.3% vs. 39% in cluster 4). Median body mass index (BMI) was the highest (34 $kg/m^2$) and HF etiology was overwhelmingly (>90%) due to nonischemic causes despite high rates of risk factors for atherosclerotic heart disease. Patients in this cluster had the highest rates of prior cerebrovascular accident and COPD, but low rates of other co-morbidities such as AF and peripheral vascular disease. They had the lowest rates of ICD and CRT use (15.7% and 4.5%), less than half of that in the next lowest group (37.4% and 19.1% in cluster 4). Cluster 2 patients were least likely to be married or employed and had the lowest levels of education and income. They exhibited objective evidence of mild HF: after cluster 4, they had the second highest median peak $VO_2$ levels (15.0 ml/kg/min) and 6MWD (351 meters). They also had the lowest median levels of NT-proBNP (418 pg/ml) and galectin-3 (11.9 ng/ml), but ST2 levels were similar to cluster 4 (21.2 vs. 21.1 ng/ml). Despite this, cluster 2 patients had the highest rates of prior hospitalization and second lowest QOL scores.

### Cluster 3 (n=313)

In terms of age, sex, and racial make-up, these patients were similar to the overall HF-ACTION study (means age: 60 years, 64% Caucasian and 75% male). HF was primarily due to ischemic cardiomyopathy (80%). The unique characteristic in these patients appeared to be their high burden of angina symptoms (97%; 43% with CCS class III or IV versus <2% for all other clusters) and, consistent with this, they had the highest rates of prior PCI and CABG. After cluster 1 patients, they had the second highest rates of ICD and CRT use. Cluster 3 patients possessed the second lowest rates of education, employment, and income. They displayed objective evidence of advanced HF, with the second lowest median peak $VO_2$ levels (14.7 ml/kg/min) and 6MWD (376 meters). Consistent with this, they had the second highest levels of all 3 prognostic biomarkers (after cluster 1): NT-proBNP (775 pg/ml), galectin-3 (14.5 ng/ml), and ST2 (23.5 ng/ml). They had the second highest rates of prior hospitalizations and the lowest QOL scores.

### Cluster 4 (n=246)

This cluster included the highest percentage of Caucasians (77%) and females (39%), with a median age of 55 years. The majority had HF due to nonischemic causes (>90%) and considerably lower rates of risk factors and comorbidities than all other patients (except for AF, which was only lower in cluster 2). Least likely to have been smokers, cluster 4 patients had the highest levels of educational attainment as well as income and were most likely to be employed at time of study onset. These patients had objective evidence of the mildest degree of HF with the highest median peak $VO_2$ levels (15.0 ml/kg/min) and 6MWD (427 meters). They also had the second lowest median levels of NTproBNP and galectin-3 (after cluster 2). These patients had the lowest rates of prior hospitalization and the highest QOL

scores. At baseline, 37.4% had an ICD and they had the second highest usage of CRT devices (19.1%).

### CLINICAL OUTCOMES

Figure 1 shows risk of primary and secondary clinical outcomes of the HF-ACTION study for each cluster, with cluster 1 (highest risk) as the comparator group. Compared with cluster 1, risk of the composite endpoint of all-cause mortality/allcause hospitalization ranged from 0.65 (0.54 to 0.78) for cluster 4 to equivalent 1.02 (0.87 to 1.19) for cluster 3. When considering all-cause mortality, cluster 3 patients demonstrated almost a 40% lower risk of mortality [0.61 (0.44 to 0.86], but risk of other outcomes was similar, suggesting a higher risk of hospitalization. Cluster 4 patients had the best risk profile, with 35% to 55% lower risk for adverse outcomes compared with cluster 1.

Figure 2 shows Kaplan-Meier curves, according to patient cluster, for the primary endpoint of all-cause death or all-cause hospitalization, and the secondary endpoint of allcause mortality. As shown, patients in clusters 1 and 3 were at the highest risk for the primary outcome, patients in cluster 4 at the lowest risk. When considering all-cause death, cluster 1 patients had the highest mortality rates, suggesting that cluster 3 patients had high rates of hospitalization.

### INTERACTION WITH EXERCISE TRAINING INTERVENTION

Benefits from exercise training, the randomized intervention tested in HF-ACTION, varied across patient clusters (Central Illustration). We found evidence of significant improvements in 3-month peak $VO_2$ levels with exercise training in cluster 2 and 3 patients: 1.33 (0.67 to 1.98) ml/min and 0.87 (0.24 to 1.51) ml/min, respectively (p for interaction = 0.04). Significant differences also were seen in the impact of exercise training on two clinical outcomes: CV death/CV hospitalization (p for interaction = 0.0396) and CV death/HF hospitalization (p for interaction = 0.0316). Clusters 1 and 2 appeared to have 12% to 30% risk reduction from exercise training whereas cluster 4 had indication of increased harm (50% to 62%); however, the confidence intervals were wide and included 1 in all cases except for the endpoint of CV death/CV hospitalization.

## DISCUSSION

We applied a novel approach to the robust database from a recent, large, randomized, controlled trial of exercise training to identify 4 clinically relevant phenotypes of chronic systolic HF. Patients within each cluster varied considerably along measures of age, sex, race, symptoms, comorbidities, HF etiology, socioeconomic status, QOL, CPET parameters, and biomarker levels. We noted differential associations with risk of hospitalization and mortality between and within clusters, as well as varied responses to exercise therapy (Central Illustration). These findings underscore the significant heterogeneity that exists within chronic HF patients and the need for improved syndrome phenotyping.

To our knowledge, this is the first application of cluster analysis to identify distinct clinical phenotypes in a large cohort of patients with chronic HF, a syndrome believed to comprise multiple disease subtypes (3). Several prior studies have used this method to successfully

identify clinically relevant patient subgroups within similarly complex, yet disparate, syndromes such as COPD, Parkinson's disease, and human encephalitis, leading to key insights about disease pathophysiology (19–22). In general, the studies' impact has been limited by their small size, low number of available clinical variables, a well-phenotyped population, and lack of outcome data. The HF-ACTION database was ideal to overcome these limitations.

The findings presented here are important for several reasons, especially when considering that LVEF—the methodology most commonly used to describe HF—was one of only a handful of variables that was statistically identical across all 4 patient clusters, emphasizing the need for improved descriptions of disease subtypes. We identified 2 clusters of patients with HF as a result of ischemic cardiomyopathy (clusters 1 and 3) that differed almost 9-fold in frequency and intensity of angina symptoms (prevalence: 11% vs. 97%; CCS angina class II to IV: 1% vs. 43%). Consequently, despite having objective measures of milder disease and much higher rates of revascularization procedures, patients in cluster 3 had much higher rates of hospitalization and the poorest QOL. Previous studies have noted the persistence of anginal symptoms in HF patients despite revascularization, suggesting that pain mechanisms in this patient population might not entirely be ameliorated by restoring epicardial blood flow (23). Despite higher rates of rehospitalization, the mortality rates for cluster 3 patients were 40% lower than cluster 1 patients. This suggests that novel strategies to improve angina symptoms in this patient subtype may be impactful (24).

We also identified a cluster of patients who tended to be young, obese African Americans, with largely nonischemic cardiomyopathy. Despite having objective evidence of milder disease based on CPET parameters and HF biomarkers, these patients had high rates of hospitalization and low QOL scores, confirming prior pre-specified analyses in this patient population (25). Intriguingly, these patients also exhibited the lowest rates of ICD use (15.7%), even though almost all qualified based on EF and NYHA criteria. Whether socioeconomic factors caused these differences is unknown, although racial and socioeconomic disparities in medical device use have been noted previously (26). Furthermore, the etiology of HF in these patients is unclear; whether it results from hypertension or other causes or it represents a distinct pathophysiological entity is an intriguing notion that requires further study (27).

Cluster 2 patients also possessed surprisingly low rates of conduction abnormalities as well as the lowest levels of biomarkers that signify myocardial stretch and fibrosis. This might explain the distinct natural history of HF previously noted in this patient population and, potentially, different responses to therapeutics (28–30). Lastly, it appears that the highest rates of rehospitalization in these patients occurred despite their objective measures of milder HF; this implies that therapies aimed at improving disease state alone would not decrease these patients' rehospitalization rates. Rather, a focused effort on understanding the global reasons for rehospitalization might result in more effective preventive methods (31,32).

The fourth cluster comprised largely Caucasian patients with the highest percentage of women (39%) and the highest socioeconomic status, as well as the mildest form of HF from

largely nonischemic cardiomyopathy. These patients had the lowest rates of comorbidities, objective measures that signified the most cardiopulmonary reserve, and the highest QOL scores. Intriguingly, exercise therapy appeared to be associated with worse outcomes in these patients. While highly speculative given the sample size, this may suggest that universal recommendations for HF patients may not always be beneficial for lower-risk patients.

Beyond what is discussed above, these data carry important implications for patient care. Whereas guidelines recommend treatment of all HF patients according to disease severity using measures that do not capture disease heterogeneity, our findings imply that it may be important to tailor therapeutics according to disease subtype based on a comprehensive evaluation of readily available clinical data. Patients resembling those in cluster 1, for example, might benefit from management of their numerous comorbid conditions along with HF, and cluster 3 patients would benefit from a focus on minimizing angina symptoms. Furthermore, the increasing use of electronic medical records may soon allow us to use clustering algorithms on large amounts of clinical data to improve phenotyping of patients and present actionable information to medical practitioners that may improve quality of care (33).

Our findings also shed light on the shortcomings of clinical trials in patients with HF: even a mechanistically sound therapeutic intervention might not show efficacy when tested on a disease state with large phenotypic variations in etiology, clinical features, and natural history (7,34). Indeed, it has been suggested that a percentage of patients in large clinical trials of HF might not even have HF, possibly explaning the high number of negative results reported in large trials of promising interventions (35). Lastly, there is a need for greater recognition of specific phenotypes within the overall HF population, which could potentially lead to targeting specific groups of patients for specific interventions.

## STUDY LIMITATIONS

Several limitations of this analysis require consideration. First and foremost, the current study was not meant to propose a new classification for chronic systolic HF, as the clusters are likely to vary according to patient characteristics and available data. These results serve to underscore the need for novel multidimensional HF classification approaches for improving patient care and trial quality. Furthermore, they are aimed to generate hypotheses for future studies that will integrate clinical and biological data on patients with the goal of improving HF phenotyping. Second, patients with incomplete datasets were excluded from cluster analyses, which necessitated complete data on individual patients. Third, the patient population represented those who participated in the HF-ACTION clinical trial and may not generalize to the entire population of chronic HF patients. Specifically, our results cannot be extrapolated to chronic HF and LVEF >35%. Fourth, the clustering algorithm yielded results based on the available clinical variables and results might have differed with more complete and accurate data. Fifth, the choice of stopping the clustering algorithm at 4 clusters included investigator discretion and preference; a larger number of clusters may refine cluster descriptions but smaller sizes may have limited our ability to explore relationships

with clinical outcomes. In summary, we considered this analysis to be hypothesis generating and further studies will be required to address these hypotheses.

## CONCLUSIONS

In conclusion, we have demonstrated that using a clustering algorithm on baseline clinical data of chronic HF patients can identify 4 phenotypically distinct and clinically meaningful groups. Patients within each cluster varied considerably along measures of age, sex, race, symptoms, comorbidities, HF etiology, socioeconomic status, QOL, CPET parameters, and biomarker levels. We also demonstrated that patients in each cluster responded distinctively to randomized intervention assignment— in this case, exercise therapy. These findings highlight the significant heterogeneity that exists within chronic HF patients and the need for improved phenotyping to enhance therapeutic efficacy.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

**Funding Sources**

Dr. Ahmad received support from the Daland Fellowship in Clinical Investigation. The HF-ACTION study was funded by grants from the National Heart, Lung, and Blood Institute (NHLBI). The biomarkers assays were funded by grants from Roche Diagnostics, BG Medicine, and Critical Diagnostics.

Drs. Felker and O'Connor have received research funding from BG Medicine, Critical Diagnostics, and Roche Diagnostics. Drs. Felker has served as a consultant for BG Medicine, Singulex, and Roche Diagnostics. Dr. Kitzman serves as a consultant for Relypsa, Inc. Dr. Piña is a consultant for General Electric and Novartis.

## Abbreviations

| | |
|---|---|
| **6MWD** | 6-minute walk distance |
| **COPD** | chronic obstructive pulmonary disease |
| **CPET** | cardiopulmonary exercise testing |
| **CV** | cardiovascular |
| **HF** | heart failure |
| **HF-ACTION** | Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training |
| **LVEF** | left ventricular ejection fraction |
| **NT-proBNP** | N-terminal pro-B-type natriuretic peptide |
| **NYHA** | New York Heart Association |

## REFERENCES

1. McMurray JJ, Adamopoulos S, Anker SD, et al. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and

Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. Eur Heart J. 2012; 33:1787–1847. [PubMed: 22611136]

2. Braunwald E. Heart Failure. JACC Heart Fail. 2013; 1:1–20. [PubMed: 24621794]
3. Loscalzo J. Personalized cardiovascular medicine and drug development: time for a new paradigm. Circulation. 2012; 125:638–645. [PubMed: 22294708]
4. De Keulenaer GW, Brutsaert DL. Systolic and diastolic heart failure: different phenotypes of the same disease? Eur J Heart Fail. 2007; 9:136–143. [PubMed: 16884955]
5. Dickstein K, Cohen-Solal A, Filippatos G, et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the Task Force for the diagnosis and treatment of acute and chronic heart failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM). Eur J Heart Fail. 2008; 10:933–989. [PubMed: 18826876]
6. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013; 62:e147–e239. [PubMed: 23747642]
7. Vaduganathan M, Greene SJ, Ambrosy AP, Gheorghiade M, Butler J. The disconnect between phase II and phase III trials of drugs for heart failure. Nat Rev Cardiol. 2013; 10:85–97. [PubMed: 23296069]
8. Butler J, Fonarow GC, Gheorghiade M. Strategies and opportunities for drug development in heart failure. JAMA. 2013; 309:1593–1594. [PubMed: 23538438]
9. National Research Council (US) Committee on A Framework for Developing a New Taxonomy of Disease. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. Washington (DC): The National Academies Press; 2011.
10. Weatherall M, Shirtcliffe P, Travers J, Beasley R. Use of cluster analysis to define COPD phenotypes. Eur Respir J. 2010; 36:472–474. [PubMed: 20930198]
11. Thum T, Galuppo P, Wolf C, et al. MicroRNAs in the human heart: a clue to fetal gene reprogramming in heart failure. Circulation. 2007; 116:258–267. [PubMed: 17606841]
12. O'Connor CM, Whellan DJ, Lee KL, et al. Efficacy and safety of exercise training in patients with chronic heart failure: HF-ACTION randomized controlled trial. JAMA. 2009; 301:1439–1450. [PubMed: 19351941]
13. Whellan DJ, O'Connor CM, Lee KL, et al. Heart Failure and A Controlled Trial Investigating Outcomes of Exercise TraiNing (HF-ACTION): Design and rationale. Am Heart J. 2007; 153:201–211. [PubMed: 17239677]
14. Hunt SA, Abraham WT, Chin MH, et al. ACC/AHA 2005 Guideline Update for the Diagnosis and Management of Chronic Heart Failure in the Adult: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Update the 2001 Guidelines for the Evaluation and Management of Heart Failure): developed in collaboration with the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: endorsed by the Heart Rhythm Society. J Am Coll Cardiol. 2005; 46:e1–e82. [PubMed: 16168273]
15. Blumenthal JA, Babyak MA, O'Connor C, et al. Effects of exercise training on depressive symptoms in patients with chronic heart failure: the HF-ACTION randomized trial. JAMA. 2012; 308:465–474. [PubMed: 22851113]
16. Felker GM, Fiuzat M, Shaw LK, et al. Galectin-3 in Ambulatory Patients with Heart Failure: Results from the HF-ACTION Study. Circ Heart Fail. 2012; 5:72–78. [PubMed: 22016505]
17. Felker GM, Fiuzat M, Thompson V, et al. Soluble ST2 in Ambulatory Patients With Heart Failure: Association With Functional Capacity and Long-Term Outcomes. Circ Heart Fail. 2013; 6:1172–1179. [PubMed: 24103327]
18. Lin D, Wei LJ, Ying Z. Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals. Biometrika. 1993; 80:557–572.
19. Weatherall M, Travers J, Shirtcliffe PM, et al. Distinct clinical phenotypes of airways disease defined by cluster analysis. Eur Respir J. 2009; 34:812–818. [PubMed: 19357143]

20. Burgel PR, Paillasseur JL, Caillaud D, et al. Clinical COPD phenotypes: a novel approach using principal component and cluster analyses. Eur Respir J. 2010; 36:531–539. [PubMed: 20075045]
21. Hamid JS, Meaney C, Crowcroft NS, Granerod J, Beyene J. Cluster analysis for identifying sub-groups and selecting potential discriminatory variables in human encephalitis. BMC Infect Dis. 2010; 10:364. [PubMed: 21192831]
22. Erro R, Vitale C, Amboni M, et al. The heterogeneity of early Parkinson's disease: a cluster analysis on newly diagnosed untreated patients. PloS One. 2013; 8:e70244. [PubMed: 23936396]
23. Mentz RJ, Fiuzat M, Shaw LK, et al. Comparison of Clinical characteristics and long-term outcomes of patients with ischemic cardiomyopathy with versus without angina pectoris (from the Duke Databank for Cardiovascular Disease). Am J Cardiol. 2012; 109:1272–1277. [PubMed: 22325975]
24. Jones DA, Timmis A, Wragg A. Novel drugs for treating angina. BMJ. 2013; 347:f4726. [PubMed: 24018101]
25. Alexander M, Grumbach K, Selby J, Brown AF, Washington E. Hospitalization for congestive heart failure. Explaining racial differences. JAMA. 1995; 274:1037–1042. [PubMed: 7563454]
26. Hernandez AF, Fonarow GC, Liang L, et al. Sex and racial differences in the use of implantable cardioverter-defibrillators among patients hospitalized with heart failure. JAMA. 2007; 298:1525–1532. [PubMed: 17911497]
27. Suthanthiran M, Li B, Song JO, et al. Transforming growth factor-beta 1 hyperexpression in African-American hypertensives: A novel mediator of hypertension and/or target organ damage. Proc Natl Acad Sci U S A. 2000; 97:3479–3484. [PubMed: 10725360]
28. Dries DL, Exner DV, Gersh BJ, Cooper HA, Carson PE, Domanski MJ. Racial differences in the outcome of left ventricular dysfunction. N Engl J Med. 1999; 340:609–616. [PubMed: 10029645]
29. Dries DL, Strong MH, Cooper RS, Drazner MH. Efficacy of angiotensin-converting enzyme inhibition in reducing progression from asymptomatic left ventricular dysfunction to symptomatic heart failure in black and white patients. J Am Coll Cardiol. 2002; 40:311–317. [PubMed: 12106937]
30. Yancy CW, Fowler MB, Colucci WS, et al. Race and the response to adrenergic blockade with carvedilol in patients with chronic heart failure. N Engl J Med. 2001; 344:1358–1365. [PubMed: 11333992]
31. Alexander M, Grumbach K, Remy L, Rowell R, Massie BM. Congestive heart failure hospitalizations and survival in California: patterns according to race/ethnicity. Am Heart J. 1999; 137:919–927. [PubMed: 10220642]
32. Mentz RJ, Bittner V, Schulte PJ, et al. Race, exercise training, and outcomes in chronic heart failure: findings from Heart Failure - a Controlled Trial Investigating Outcomes in Exercise TraiNing (HF-ACTION). Am Heart J. 2013; 166:488–495. [PubMed: 24016498]
33. Jha AK. The promise of electronic records: around the corner or down the road? JAMA. 2011; 306:880–881. [PubMed: 21862751]
34. Felker GM, Pang PS, Adams KF, et al. Clinical trials of pharmacological therapies in acute heart failure syndromes: lessons learned and directions forward. Circ Heart Fail. 2010; 3:314–325. [PubMed: 20233993]
35. Shah SJ, Heitner JF, Sweitzer NK, et al. Baseline characteristics of patients in the treatment of preserved cardiac function heart failure with an aldosterone antagonist trial. Circ Heart Fail. 2013; 6:184–192. [PubMed: 23258572]

## PERSPECTIVES

### Competency in Medical Knowledge

There is considerable heterogeneity among patients with chronic heart failure related to etiology, clinical manifestations, and natural history; certain characteristics identified by cluster analysis are associated with differences in outcomes.

### Competency in Patient Care

In managing patients with chronic heart failure, therapy should be individualized based on recognition of key clinical characteristics.

### Translational Outlook

To improve outcomes, clinical trials should evaluate responses to specific therapeutic interventions in defined subgroups of patients with chronic heart failure distinguished by cluster analysis.

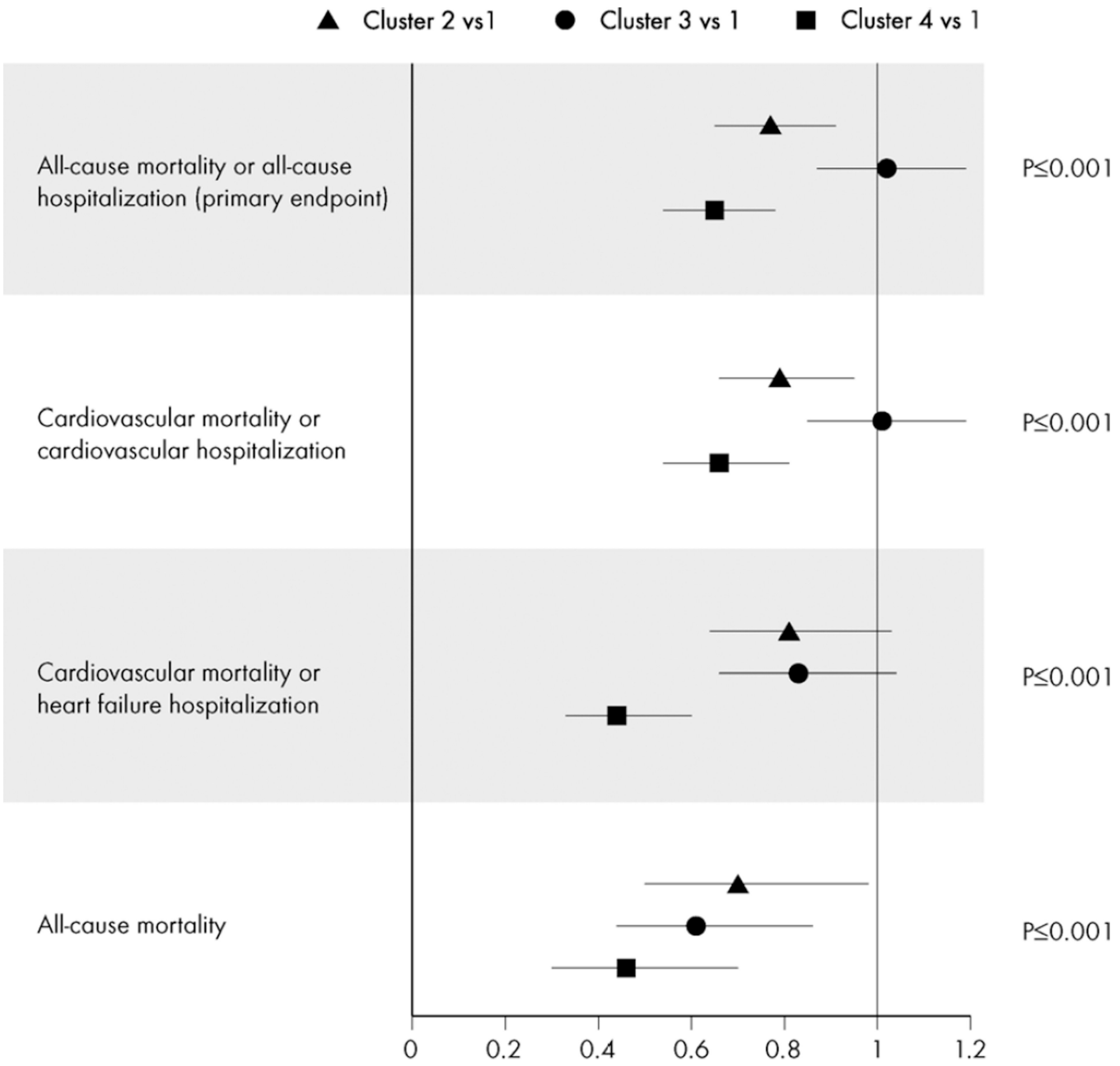
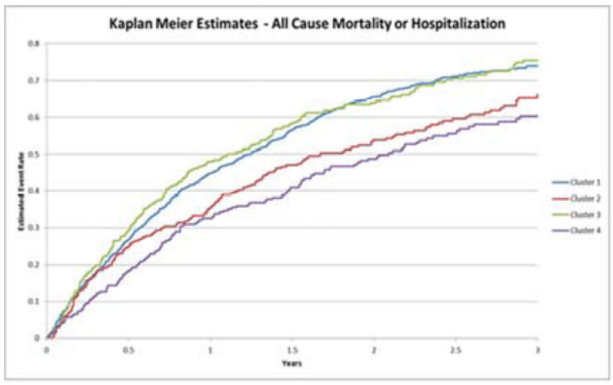



**Figure 1. Risk of Clinical Events Compared with Cluster 1 (Highest Risk)***
Symbols represent hazard ratios and 95 % confidence intervals.



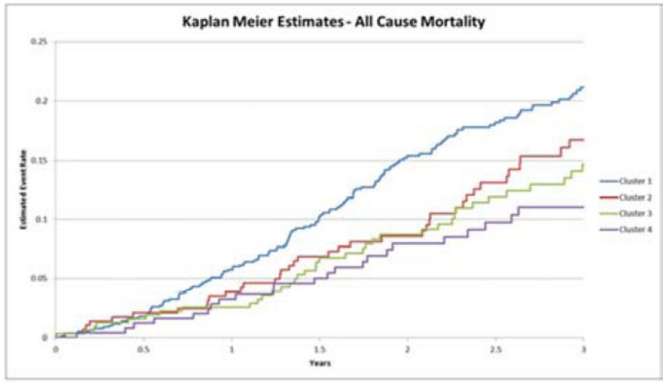
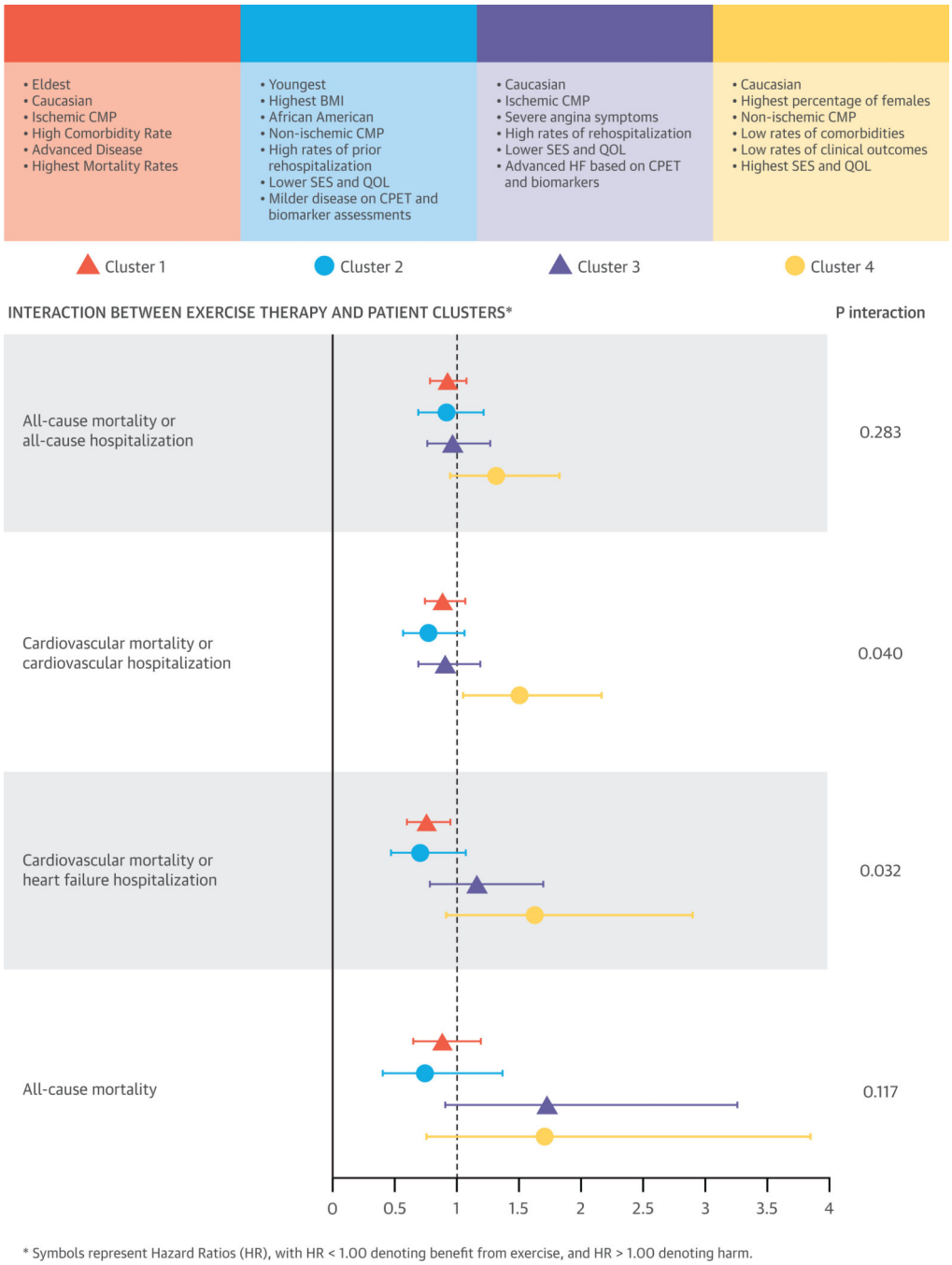




**Figure 2. Time to All-Cause Mortality/All-Cause Hospitalization and to All-Cause Mortality**

Kaplan-Meier curves, according to patient cluster, that depict (A) the primary endpoint of death or hospitalization (from any cause), and (B) the secondary endpoint of death from any cause.




**Central Illustration. Consort diagram showing cluster methodology applied to HF-ACTION study and the four distinct clusters that emerge**

Interaction between exercise therapy and patient clusters extracted from the study. Symbols represent Hazard Ratios (HR) and 95 % Confidence Intervals (CI), with HR < 1.00 denoting benefit from exercise, and HR > 1.00 denoting harm.

*J Am Coll Cardiol*. Author manuscript; available in PMC 2015 October 28.

**TABLE 1**

Baseline Clinical Characteristics According to Patient Clusters

| Characteristic | Cluster 1 (n = 773) | Cluster 2 (n = 287) | Cluster 3 (n = 313) | Cluster 4 (n = 246) | p Value* |
|---|---|---|---|---|---|
| Age, yrs | 63(56–72) | 49(40–56) | 60(53–68) | 55(46–64) | <0.001 |
| Female, % | 21 | 38 | 25 | 39 | <0.001 |
| Black, % | 28 | 69 | 28 | 20 | <0.001 |
| White, % | 67 | 27 | 67 | 77 | <0.001 |
| BMI, kg/m$^2$ | 30(26–34) | 34(27–41) | 29(27–33) | 28(25–33) | <0.001 |
| Systolic BP, mm Hg | 112(102–126) | 117(108–130) | 114(102–130) | 104(94–114) | <0.001 |
| Diastolic BP, mm Hg | 70(62–78) | 76(68–84) | 70(60–80) | 64(60–72) | <0.001 |
| Ischemic cardiomyopathy, % | 68 | 10 | 80 | 9 | <0.001 |
| **Prior heart failure hospitalizations** | | | | | |
| None, % | 76.8 | 56.1 | 74.4 | 81.3 | <0.001 |
| 1, % | 17.9 | 33.4 | 19.2 | 14.6 | |
| 2, % | 3.5 | 7.7 | 3.2 | 2.4 | |
| ≥3, % | 1.8 | 2.8 | 3.2 | 1.6 | |
| **Symptoms** | | | | | |
| NYHA III–IV, % | 39 | 27 | 43 | 21 | <0.001 |
| History of angina, % | 11 | 14 | 97 | 7 | <0.001 |
| CCS Angina Class, % | | | | | <0.001 |
| 0 | 95 | 90 | 29 | 98 | |
| 1 | 4 | 8 | 28 | 2 | |
| 2–4 | 1 | 2 | 43 | 0 | |
| **Past medical and surgical history** | | | | | |
| History of MI, % | 55.2 | 6.6 | 70.9 | 6.1 | <0.001 |
| Hypertension, % | 71.2 | 64.5 | 74.1 | 12.6 | <0.001 |
| Diabetes, % | 41.9 | 21.6 | 41.2 | 5.7 | <0.001 |
| Atrial fibrillation/flutter, % | 31.8 | 6.6 | 14.7 | 11.8 | <0.001 |
| Hyperlipidemia, % | 76.8 | 40.8 | 83.7 | 38.6 | <0.001 |

*J Am Coll Cardiol*. Author manuscript; available in PMC 2015 October 28.

| Characteristic | Cluster 1 (n = 773) | Cluster 2 (n = 287) | Cluster 3 (n = 313) | Cluster 4 (n = 246) | p Value* |
|---|---|---|---|---|---|
| Stroke, % | 12.0 | 12.5 | 9.9 | 4.1 | 0.003 |
| PVD, % | 9.7 | 4.2 | 6.7 | 0.8 | <0.001 |
| COPD, % | 12.7 | 13.2 | 8.9 | 4.5 | 0.001 |
| Prior valve surgery, % | 8.0 | 1.0 | 3.2 | 4.9 | <0.001 |
| Prior PCI, % | 27.8 | 3.1 | 42.8 | 2.8 | <0.001 |
| Prior CABG, % | 37.1 | 1.7 | 42.2 | 1.6 | <0.001 |
| **Laboratories** | | | | | |
| Sodium, mmol/l | 139(137–141) | 139(138–141) | 139(137–141) | 139(137–140) | 0.033 |
| Creatinine, mg/dl | 1.3(1.1–1.6) | 1.1(0.9–1.3) | 1.2(1.0–1.4) | 1.1(0.9–1.2) | <0.001 |
| Blood urea nitrogen, mg/dl | 23(17–32) | 16(12–20) | 20(15–26) | 18(14–24) | <0.001 |
| **Medications and devices** | | | | | |
| ACE-I or ARB, % | 93.4 | 94.8 | 93.3 | 97.2 | 0.140 |
| Beta-blocker, % | 95.0 | 95.8 | 95.2 | 93.9 | 0.787 |
| Loop diuretic, % | 81.0 | 81.9 | 78.0 | 69.5 | <0.001 |
| Digoxin, % | 49.5 | 43.2 | 43.5 | 49.6 | 0.118 |
| ICD, % | 53.3 | 15.7 | 39.6 | 37.4 | <0.001 |
| CRT, % | 25.2 | 4.5 | 14.7 | 19.1 | <0.001 |
| **Resting ECG conduction** | | | | | <0.001 |
| Normal, % | 31.8 | 73.9 | 46.6 | 37.0 | |
| LBBB, % | 14.9 | 8.7 | 16.3 | 26.4 | |
| RBBB, % | 4.8 | 2.8 | 4.2 | 1.6 | |
| IVCD, % | 14.9 | 10.8 | 13.7 | 14.2 | |
| Paced, % | 33.6 | 3.8 | 19.2 | 20.7 | |

Values are median(interquartile range), or %.

*p values for the comparisons of variables across clusters.

ACE–I = angiotensin–converting enzyme inhibitor; ARB = angiotensin receptor blocker; BMI = body mass index; BP = blood pressure; CABG = coronary artery bypass grafting; CCS = Canadian Cardiovascular Society; COPD = chronic obstructive pulmonary disease; CRT = cardiac resynchronization therapy; ECG = electrocardiogram; HF, heart failure; ICD = implantable cardioverter–defibrillator; IVCD = intraventricular conduction delay; LBBB = left bundle branch block; MI = myocardial infarction; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; PVD= peripheral vascular disease; RBBB = right bundle branch block.

**TABLE 2**

Baseline Psychosocial Characteristics According to Patient Clusters

| Characteristic | Cluster 1 (n = 773) | Cluster 2 (n = 287) | Cluster 3 (n = 313) | Cluster 4 (n = 246) | p Value* |
|---|---|---|---|---|---|
| **Marital status** | | | | | |
| Married, % | 64 | 38 | 63 | 60 | <0.001† |
| Divorced, % | 13 | 16 | 14 | 14 | |
| Single(Never Married), % | 7 | 25 | 8 | 13 | |
| Other, % | 16 | 21 | 15 | 13 | |
| **Smoking status** | | | | | |
| Never, % | 31 | 44 | 32 | 59 | <0.001 |
| Current, % | 14 | 28 | 13 | 9 | |
| Past, % | 54 | 28 | 55 | 35 | |
| **Alcohol use, %** | 43 | 43 | 46 | 43 | 0.864 |
| **Highest level of education** | | | | | |
| Less than high school, % | 11 | 12 | 14 | 5 | 0.004† |
| High school, % | 28 | 32 | 29 | 25 | |
| Associate degree, % | 9 | 11 | 10 | 10 | |
| College, % | 19 | 9 | 12 | 22 | |
| Graduate school, % | 9 | 5 | 6 | 16 | |
| Other, % | 24 | 31 | 29 | 22 | |
| **Employment Status** | | | | | |
| Employed full time, % | 16 | 23 | 14 | 27 | <0.001† |
| Employed part time, % | 6 | 5 | 6 | 5 | |
| Disabled, % | 27 | 44 | 38 | 28 | |
| Unemployed, % | 4 | 13 | 4 | 7 | |
| Retired, % | 47 | 11 | 34 | 28 | |
| Other, % | 0 | 4 | 4 | 5 | |
| **Income** | | | | | |
| <d$15,000, % | 17 | 27 | 24 | 13 | <0.001† |

| Characteristic | Cluster 1 (n = 773) | Cluster 2 (n = 287) | Cluster 3 (n = 313) | Cluster 4 (n = 246) | p Value* |
|---|---|---|---|---|---|
| $15,000 –$24,999, % | 16 | 18 | 17 | 17 | |
| $25,000 –$34,999, % | 13 | 14 | 16 | 12 | |
| $35,000 – $49,999, % | 16 | 13 | 15 | 11 | |
| $50,000 – $74,999, % | 15 | 10 | 14 | 19 | |
| $75,000 – $99,999, % | 8 | 4 | 6 | 8 | |
| >$100,000, % | 6 | 3 | 4 | 12 | |
| **Quality of Life** | | | | | |
| KCCQ Score | 72(54–85) | 63(43–80) | 60(47–76) | 76(60–86) | <0.001 |
| BDI-II Score | 8(4–13) | 10(5–19) | 10(6–16) | 7(4–13) | <0.001 |
| Euro Thermometer | 70(60–80) | 66(50–80) | 65(50–80) | 70(60–80) | <0.001 |

Values are median(interquartile range) or %.

*p values for comparisons across clusters.

†p values for the dichotomized comparison of each variable as follows: income: <$25,000 vs. ≥$25,000; education: <high school vs. ≥high school; marital status: positive current or prior partner(married, living with partner, widowed) vs. no partner(single, divorced, separated); employment status: employed, volunteer, student, homemaker, or retired vs. unemployed or disabled.

BDI–II = Beck Depression Inventory-II; KCCQ = Kansas City Cardiomyopathy Questionnaire.

**TABLE 3**

Objective Predictors of Heart Failure Prognosis According to Patient Clusters

| Patient Biomarkers | Cluster 1 (n = 773) | Cluster 2 (n = 287) | Cluster 3 (n = 313) | Cluster 4 (n = 246) | p Value |
|---|---|---|---|---|---|
| LVEF, % | 25(20–30) | 25(20–30) | 25(20–30) | 24(19–30) | 0.606 |
| Peak $VO_2$, ml/kg/min | 13.5(11.0–16.5) | 15.0(12.1–18.0) | 14.7(12.0–17.9) | 17.5(14.2–20.7) | <0.001 |
| $VEVCO_2$ slope | 34(30–40) | 30(26–34) | 33(29–39) | 31(27–35) | <0.001 |
| 6MWD, meters | 351(290–416) | 394(320–439) | 376(305–441) | 427(363–476) | <0.001 |
| NT-proBNP, pg/ml(n = 1,011 ) | 1079(461–2517) | 418(194–978) | 775(359–1663) | 558(206–1606) | <0.001 |
| Galectin-3, ng/ml(n = 664) | 15.4(11.9–21.0) | 11.9(9.8–14.9) | 14.5(10.8–20.1) | 12.3(10.2–16.7) | <0.001 |
| ST2, ng/ml(n = 677) | 26.2(20.5–35.1) | 21.2(15.7–28.3) | 23.5(19.0–30.5) | 21.1(16.3–26.7) | <0.001 |

Values are median(interquartile range).

6MWD = 6-minute walk distance; LVEF = left ventricular ejection fraction; NT-proBNP = Nterminal pro-B-type natriuretic peptide; $VO_2$ = peak oxygen consumption per unit of time; $VEVCO_2$ = minute ventilation - carbon dioxide production relationship.