ORIGINAL RESEARCH

Discordance Between Severity of Heart Failure as Determined by Patient Report Versus Cardiopulmonary Exercise Testing

Katherine C. Michelis, MD; Justin L. Grodin [iD], MD, MPH; Lin Zhong, MPH; Ambarish Pandey [iD], MD, MSCS; Kathleen Toto, RN, MSN, ACNP-C; Colby R. Ayers, MS; Jennifer T. Thibodeau [iD], MD, MSCS; Mark H. Drazner [iD], MD, MSc

BACKGROUND: Patient-reported outcomes may be discordant to severity of illness as assessed by objective parameters. The frequency of this discordance and its influence on clinical outcomes in patients with heart failure is unknown.

METHODS AND RESULTS: In HF-ACTION (Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training), participants (N=2062) had baseline assessment of health-related quality of life via the Kansas City Cardiomyopathy Clinical Summary score (KCCQ-CS) and objective severity by cardiopulmonary stress testing (minute ventilation [V_E]/carbon dioxide production [VCO_2] slope). We defined 4 groups by median values: 2 concordant (lower severity: high KCCQ-CS and low V_E/VCO_2 slope; higher severity: low KCCQ-CS and high V_E/VCO_2 slope) and 2 discordant (symptom minimizer: high KCCQ-CS and high V_E/VCO_2 slope; symptom magnifier: low KCCQ-CS and low V_E/VCO_2 slope). The association of group assignment with mortality was assessed in adjusted Cox models. Symptom magnification (23%) and symptom minimization (23%) were common. Despite comparable KCCQ-CS scores, the risk of all-cause mortality in symptom minimizers versus concordant–lower severity participants was increased significantly (hazard ratio [HR], 1.79; 95% CI, 1.27–2.50; P<0.001). Furthermore, despite symptom magnifiers having a KCCQ-CS score 28 points lower (poorer QOL) than symptom minimizers, their risk of mortality was not increased (HR, 0.79; 95% CI, 0.57–1.1; P=0.18, respectively).

CONCLUSIONS: Severity of illness by patient report versus cardiopulmonary exercise testing was frequently discordant. Mortality tracked more closely with the objective data, highlighting the importance of relying not only on patient report, but also objective data when risk stratifying patients with heart failure.

Key Words: dyspnea ■ epidemiology ■ prognosis ■ quality of life ■ stress test

Heart failure is associated with significant morbidity and mortality despite advances in diagnosis and management.[1–3] Risk stratification is an integral aspect of caring for patients with heart failure, allowing alignment of intensity of therapeutic interventions to severity of illness. Recently, there has been increasing interest in patient-reported outcomes (PROs) as assessed by validated health-related quality of life (HR-QOL) instruments such as the Kansas City Cardiomyopathy Questionnaire (KCCQ).[4] PROs

are being incorporated as end points in cardiovascular clinical trials and as performance metrics in registries, and they also are now being evaluated to support product labeling by the US Food and Drug Administration.[5] Thus, a full understanding of the strengths and limitations of PROs is increasingly important.

In our clinical practice, we have encountered patients whose reported symptoms appear discordant to objective markers of heart failure severity of illness. In some cases, individuals with marked abnormalities

Nonstandard Abbreviations and Acronyms

HF-ACTION	Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training
KCCQ	Kansas City Cardiomyopathy Questionnaire
KCCQ-CS	Kansas City Cardiomyopathy Questionnaire clinical summary score
NYHA	New York Heart Association
PRO	patient-reported outcome
VCO₂	carbon dioxide production
V$_E$	minute ventilation
VO₂	oxygen uptake

in objective measures of heart failure severity, such as natriuretic peptide levels, echocardiography, invasive hemodynamic assessment, or cardiopulmonary stress testing, report minimal symptoms; whereas in other cases, individuals with less severe objective data report severe symptoms. Although prior studies have assessed the discordance between patient-reported HR-QOL and objective measures of heart failure severity,[6–11] clinical factors associated with this discordance and its implications for prognosis have not been fully elucidated.

To address these gaps in knowledge, we queried the HF-ACTION (Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training) database. HF-ACTION provided an ideal opportunity to address these questions, because participants were carefully phenotyped for severity of illness both via cardiopulmonary exercise testing and by patient-reported symptoms ascertained by the KCCQ, and then followed for clinical outcomes over several years.[12,13]

METHODS

Study Setting

This study is a post hoc analysis of HF-ACTION. This article was prepared using research materials provided by the National Heart, Lung, and Blood Institute's Biologic Specimen and Data Repository Information Coordinating Center via an approved proposal. Other investigators may request to access the data set at https://biolincc.nhlbi.nih.gov/studies/hf_action/.

A complete description of the trial's design and major findings have been published previously.[12–14] In brief, HF-ACTION was a multicenter, randomized controlled trial that compared the safety and efficacy of exercise training and medical therapy versus medical therapy alone in patients with heart failure with reduced ejection fraction (left ventricular ejection fraction <35%) and New York Heart Association class II–IV symptoms despite optimal medical therapy. Between April 2003 and February 2007, a total of 2331 patients from 82 centers within the United States, Canada, and France were randomized.

For our primary analysis, the KCCQ clinical summary score (KCCQ-CS) was used as the subjective measure of heart failure severity, and ventilatory efficiency, as assessed by the slope of minute ventilation (V_E) versus carbon dioxide production (VCO_2), was used as the objective measure (V_E/VCO_2 slope). Participants were included only if they had values for both measures at baseline. Median follow-up was 441 days (interquartile range, 160–802 days). A priori, we elected to focus on KCCQ-CS rather than the overall summary score, because the KCCQ-CS is more restrictive than the overall summary score sinceit includes only the total symptom and physical function scores, domains that may be more likely to correlate with objective disease severity as assessed by cardiopulmonary exercise testing.

V_E/VCO_2 slope describes the rate of increase of ventilation in relation to CO_2 production during exercise, and in patients with heart failure, an excessive ventilatory response is reflected by an elevated V_E/VCO_2 slope.[15] Although both peak oxygen uptake (VO_2) and V_E/VCO_2 slope are well-accepted measures of heart failure severity used to risk stratify patients for consideration of heart transplantation,[16] V_E/VCO_2 slope provides independent prognostic information beyond peak VO_2.[17–19] Moreover, V_E/VCO_2 slope is accurate despite submaximal effort during cardiopulmonary exercise testing[20] and therefore should not be influenced by psychological factors that may limit patient motivation. Although VO_2 at the anaerobic threshold is another prognostic variable that could be evaluated despite submaximal effort on cardiopulmonary exercise testing, this value is not always able to be determined, especially for patients with more advanced heart failure.[21] Therefore, we selected V_E/VCO_2 slope as the objective measure of heart failure severity. However, we also performed confirmatory analyses using peak VO_2 in a subgroup restricted to respiratory exchange ratio ≥1.1, a value that has been used to demonstrate that the patient has performed with maximal effort.[22,23]

Concordant and Discordant Group Assignment

The study cohort was classified into 4 groups on the basis of the median values of KCCQ-CS and V_E/VCO_2 slope (Figure 1). A higher score on the KCCQ-CS indicates better HR-QOL,[24] and lower V_E/VCO_2 slope reflects lower severity of illness.[25] Therefore, concordant groups consisted of participants with high KCCQ-CS and low V_E/VCO_2 slope (concordant–lower severity) or with low KCCQ-CS and high V_E/VCO_2 slope (concordant–higher severity). Discordant groups consisted of participants with low KCCQ-CS and low V_E/VCO_2 slope (symptom magnifiers) or with high KCCQ-CS and high V_E/VCO_2 slope (symptom minimizers). For the secondary analysis based on peak VO_2, a similar method of classification was performed using median peak VO_2, with a higher peak VO_2 reflecting lower severity of illness.[26]

Clinical Outcomes

The primary composite end point of the HF-ACTION trial was all-cause mortality or all-cause hospitalization. Additional clinical outcomes considered in this analysis included the individual components of the
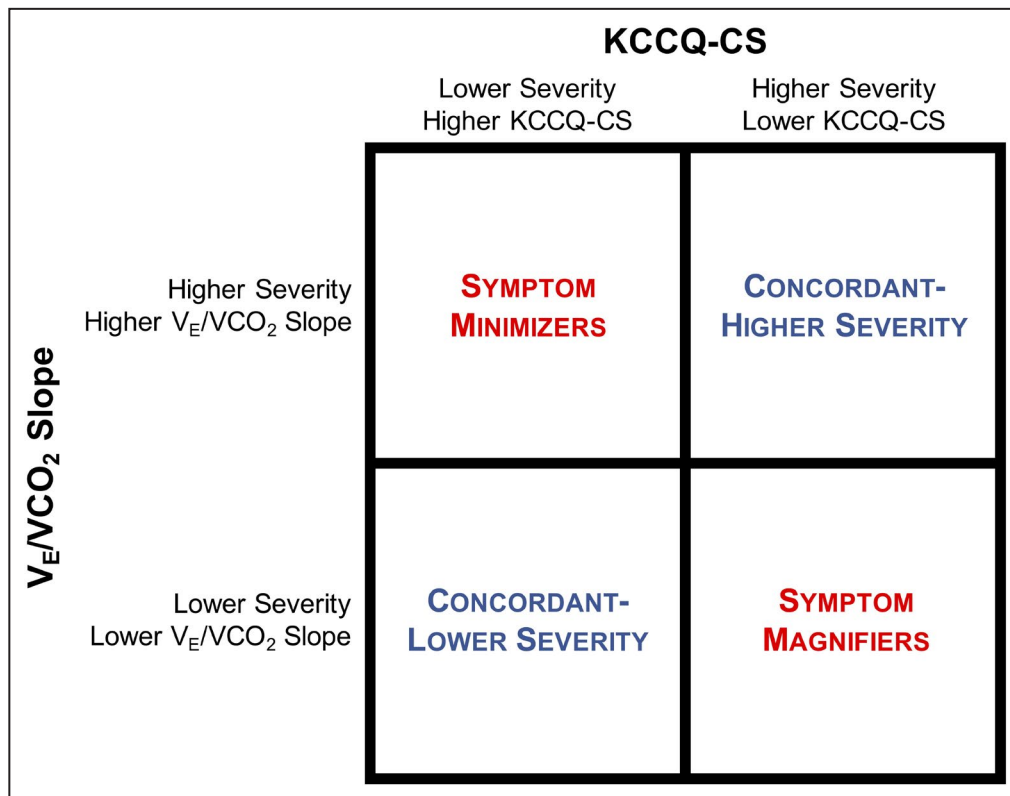

Figure 1. Defining group classification by KCCQ and V_E/VCO_2 slope.
This diagram demonstrates the categorization of the study cohort based on median values of KCCQ-CS and V_E/VCO_2 slope. KCCQ-CS indicates Kansas City Cardiomyopathy Questionnaire Clinical Summary score; and V_E/VCO_2, slope of minute ventilation/carbon dioxide production.

composite outcome (all-cause mortality and all-cause hospitalization), the Beck Depression Inventory (BDI) II score, Multidimensional Scale of Perceived Social Support, Borg rating at peak exercise, and EuroQOL 5-dimension questionnaire. A higher BDI II score indicates more severe depression.[27] Higher scores on the Multidimensional Scale of Perceived Social Support indicate better perceived availability and sufficiency of social support,[28] and a higher Borg rating indicates greater perceived exertion during physical activity.[29] The EuroQOL 5-dimension questionnaire is a self-administered tool that has been validated in patients with heart failure and comprises a questionnaire and visual analog self-rating scale.[30]

Statistical Analysis

Univariate analyses comparing patient demographic and clinical factors among all 4 concordant and discordant groups and between pairs of these groups were performed. Continuous variables were expressed as median with interquartile range and compared using the Kruskal-Wallis test. Categorical variables were expressed as number (percent) and compared using the χ^2 test. Linear regression was used to test the association between KCCQ-CS and V_E/VCO_2 slope, and the coefficient of determination (R^2) was reported. Additionally, multivariable linear regression, adjusted for age, sex, and race, was performed to examine the associations of concordant and discordant group assignment with BDI II score, Multidimensional Scale of Perceived Social Support, and Borg rating at peak exercise. Multivariable logistic regression was performed, adjusting for age, sex, and race, to examine the associations of concordant and discordant group assignment with specific questions on the BDI II and EuroQOL 5-dimension questionnaire. Variables were dichotomized as follows on the basis of participant responses: 0 or 1 to 3 for the BDI II questionnaire and no limiting symptoms or any limiting symptoms on the EuroQOL 5-dimension questionnaire.

Kaplan-Meier estimates of time-to-event clinical outcomes were calculated, and survival curves were compared using the log-rank test. Cox proportional hazards models of these study outcomes were performed using concordant and discordant group assignment as the primary exposure, adjusting for age, sex, and race in each model, with the concordant–lower severity group as the reference. In additional multivariable analyses, we adjusted for depression and obesity. A pairwise comparison between the symptom minimizer and symptom magnifier groups for risk of mortality was also performed. The proportional hazards assumption for each of these models was validated by Schoenfeld residuals.

Two-sided $P<0.05$ were considered significant in this study. Statistical analyses were completed with R version 3.6.0 (R Foundation for Statistical Computing, Vienna, Austria) or SAS version 9.4 software (SAS Institute, Cary, NC).

RESULTS

Frequency of Discordance Within the Study Population

A total of 2062 participants had available KCCQ-CS and V_E/VCO_2 slope data at baseline and were included in this study. KCCQ-CS and V_E/VCO_2 slope were weakly correlated ($R^2=0.014$, $P<0.001$; Figure S1). When participants were stratified above or below the median for heart failure severity of illness by patient report (KCCQ-CS) and cardiopulmonary exercise testing (V_E/VCO_2 slope), the concordance rate for both subjective and objective measures was only slightly better than chance (54.1%).

Patient Characteristics Associated With Discordance

Table 1 displays the baseline characteristics stratified by the 4 groups. Relative to the concordant–lower severity participants, symptom magnifiers were more likely to be classified as New York Heart Association (NYHA) class III, be depressed, and have a lower HR-QOL as assessed by other PRO measurements like the EuroQoL Visual Analog Scale ($P<0.001$ for all comparisons). Symptom magnifiers were also more likely to have diabetes mellitus, chronic obstructive pulmonary disease, and obesity than the concordant–lower severity group. Of note, symptom magnifiers versus the concordant–lower severity group had lower scores on the KCCQ-CS and EuroQoL Visual Analog Scale by ≥30 and ≥16 points, respectively. Symptom minimizers were less likely to be depressed, be obese or NYHA class III, and had lower BDI-II scores and higher EuroQoL Visual Analog Scale scores, as compared with the concordant–higher severity group. The blood urea nitrogen, creatinine, and B-type natriuretic peptide levels were not different between the symptom magnifiers versus concordant–lower severity group and between the symptom minimizers versus concordant–higher severity group.

We next compared other parameters obtained from the cardiopulmonary exercise testing between these groups (Table 2). Symptom magnifiers, versus those in the concordant–lower severity group, had lower 6-minute-walk distances, cardiopulmonary exercise test duration, peak VO_2, and VO_2 at ventilatory threshold, despite similar VE/VCO_2 slopes. In

Table 1. Baseline Characteristics Stratified by Concordant and Discordant Group Assignment

| | Favorable V_E/VCO_2 Slope | | Unfavorable V_E/VCO_2 Slope | | |
Characteristic	High KCCQ-CS, Concordant–Lower Severity, N=566, 27.4%	Low KCCQ-CS, Discordant: Symptom Magnifier, N=480, 23.3%	High KCCQ-CS, Discordant: Symptom Minimizer, N=465, 22.6%	Low KCCQ-CS, Concordant–Higher Severity, N=551, 26.7%	P Value*
Age, y	57 (48–66)	55 (46–62)[†,‡]	64 (56–73)[§]	60 (53–70)	<0.001
Men	407 (71.9%)	338 (70.4%)	342 (73.5%)	396 (71.9%)	0.77
Race					0.065
Black	184 (32.5%)	167 (34.8%)[‖]	126 (27.1%)	179 (32.5%)	
White	345 (61.0%)	275 (57.3%)[‖]	315 (67.7%)	336 (61.0%)	
Other	37 (6.5%)	38 (7.9%)[‖]	24 (5.2%)	36 (6.5%)	
NYHA class					<0.001
II	474 (83.7%)	265 (55.2%)[‡,¶]	341 (73.3%)[§]	237 (43.0%)	
III	92 (16.3%)	215 (44.8%)[‡,¶]	124 (26.7%)[§]	314 (57.0%)	
Ischemic cause of heart failure	236 (41.7%)	207 (43.1%)[‡]	302 (64.9%)[#]	312 (56.6%)	<0.001
Left ventricular ejection fraction, %, N=1911	27 (22–32)	25 (20–30)[†,‖]	24 (19–29)	23 (20–29)	<0.001
Diabetes mellitus	145 (25.6%)	172 (35.8%)[†]	154 (33.1%)	192 (34.8%)	0.001
Chronic obstructive pulmonary disease, N=2043	44 (7.9%)	58 (12.2%)[†]	44 (9.6%)[#]	77 (14.0%)	0.006
Obesity, N=2058	286 (50.6%)	317 (66.2%)[‡,¶]	151 (32.5%)[§]	256 (46.6%)	<0.001
Depression	100 (17.7%)	146 (30.4%)[‡,¶]	66 (14.2%)[§]	129 (23.4%)	<0.001
KCCQ-CS	88 (81–94)	58 (47–68)	86 (80–94)	58 (47–66)	N/A
EuroQoL Visual Analog Scale score, N=2012	76 (65–86)	60 (45–70)[‡,¶]	75 (60–80)[§]	60 (40–70)	<0.001
Beck Depression Inventory II score, N=2055	6 (3–9)	13 (8–20)[‡,¶]	5 (3–8)[§]	12 (7–20)	<0.001
Multidimensional Scale of Perceived Social Support, N=2056	6.2 (5.4–6.8)	5.7 (4.8–6.5)[‡,¶]	6.2 (5.6–6.8)[§]	5.8 (5.0–6.7)	<0.001
Body mass index, N=2058	30 (27–36)	33 (28–39)[‡,¶]	28 (24–31)[§]	29 (25–34)	<0.001
Systolic blood pressure, mm Hg, N=2058	116 (105–130)	112 (102–124)[†]	110 (100–127)	110 (100–121)	<0.001
Blood urea nitrogen, mg/dL, N=1793	19 (14–24)	18 (14–25)[‡]	22 (17–31)	23 (17–32)	<0.001
Creatinine, mg/dL, N=1850	1.1 (1.0–1.3)	1.1 (1.0–1.4)[‡]	1.2 (1.0–1.6)	1.3 (1.0–1.6)	<0.001
B-type natriuretic peptide, pg/mL, N=723	142 (69–323)	148 (78–364)[‡]	396 (209–795)	311 (161–661)	<0.001
Medications					
ß-blocker	535 (94.5%)	447 (93.1%)	444 (95.5%)	521 (94.6%)	0.46
ACE inhibitor or ARB	534 (94.3%)	453 (94.4%)	443 (95.3%)	513 (93.1%)	0.52
Aldosterone receptor antagonist	223 (39.4%)	231 (48.1%)[†]	194 (41.7%)[#]	268 (48.6%)	0.003

Values are presented as median (interquartile range) for continuous variables and number (percent) for categorical variables. Because of missing data, N is listed for those variables where the cohort with available data was less than N=2062. Obesity was defined as body mass index ≥30 kg/m². ACE indicates angiotensin converting enzyme; ARB, angiotensin II receptor blocker; EuroQoL, European Quality of Life; KCCQ-CS, Kansas City Cardiomyopathy Questionnaire Clinical Summary score; NYHA, New York Heart Association; and V_E/VCO_2, minute ventilation/carbon dioxide production.

*P for comparison among all 4 groups.

†P<0.05 vs concordant–lower severity group.

‡P<0.001 vs symptom minimizer group.

§P<0.001 vs concordant–higher severity group.

‖P<0.05 vs symptom minimizer group.

¶P<0.001 vs concordant–lower severity group.

#P<0.05 vs concordant–higher severity group.

Table 2. Functional Status Measures at Baseline Stratified by Concordant and Discordant Group Assignment

Characteristic	Favorable V_E/VCO_2 Slope		Unfavorable V_E/VCO_2 Slope		P Value*
	High KCCQ-CS, Concordant–Lower Severity, N=566, 27.4%	Low KCCQ-CS, Discordant: Symptom Magnifier, N=480, 23.3%	High KCCQ-CS, Discordant: Symptom Minimizer, N=465, 22.6%	Low KCCQ-CS, Concordant–Higher Severity, N=551, 26.7%	
Distance of 6 min walk, m, N=2016	411 (350–463)	367 (297–442)†	377 (318–430)‡	323 (251–394)	<0.001
Cardiopulmonary exercise duration, min, N=2044	12 (9–14)	10 (7–12)†,§	9 (7–11)‡	8 (6–10)	<0.001
Borg rating at peak exercise, N=1998	17 (15–18)	17 (15–19)§,‖	17 (15–19)¶	17 (15–19)	<0.001
V_E/VCO_2	28 (26–31)	28 (26–31)	38 (35–43)	39 (35–46)	N/A
Peak VO_2, mL/kg per min, N=2061	17 (14–20)	15 (13–19)†,#	13 (11–16)‡	12 (10–15)	<0.001
Oxygen uptake efficiency slope, N=2053	2.0 (1.6–2.4)	2.0 (1.6–2.4)#	1.3 (1.0–1.6)	1.3 (1.0–1.7)	<0.001
VO_2 at the ventilatory threshold, mL/kg per min, N=1771	12 (10–14)	11 (9–13)†,#	10 (9–12)¶	10 (8–11)	<0.001

Values are presented as median (interquartile range) for continuous variables and number (percent) for categorical variables. Because of missing data, N is listed for those variables where the cohort with available data was less than N=2062. KCCQ-CS indicates Kansas City Cardiomyopathy Questionnaire Clinical Summary score; V_E/VCO_2, minute ventilation/carbon dioxide production; and VO_2, oxygen uptake.

*P for comparison among all 4 groups.

†P<0.001 vs concordant–lower severity group.

‡P<0.001 vs concordant–higher severity group.

§P<0.05 vs symptom minimizer group.

‖P<0.05 vs concordant–lower severity group.

¶P<0.05 vs concordant–higher severity group.

#P<0.001 vs symptom minimizer group.

comparison to concordant–higher severity participants, symptom minimizers, despite a similar VE/VCO2, had more favorable functional data overall with higher 6-minute-walk distances, cardiopulmonary exercise test duration, peak VO_2, and VO_2 at ventilatory threshold.

Characteristics of Symptom Minimizers Versus Symptom Magnifiers

Relative to symptom minimizers, those in the symptom magnifier group had lower B-type natriuretic peptide levels (P<0.001) and had higher peak VO_2, oxygen uptake efficiency slope, and VO_2 at ventilatory threshold (P<0.001 for all), although their KCCQ-CS was 28 points lower (Tables 1 and 2).

Association Between Discordance and Patient Characteristics in Adjusted Models

In adjusted models, symptom magnifier status, as compared with concordant–lower severity status, was associated with more perceived limitations, as assessed by higher BDI II score (ß=7.34, P<0.001), lower Multidimensional Scale

of Perceived Social Support score (ß=−0.46, P<0.001), and higher Borg rating (ß=0.38, P=0.012) (Table S1). Similar associations, in comparison to the concordant–lower severity group, were seen in the concordant–higher severity group, but not the symptom minimizer group. When we focused on answers to questions that may reflect a tendency for negative thinking, symptom magnifiers were more likely to report pessimism (odds ratio [OR], 1.3; 95% CI, 1.2–1.3; P<0.001), past failure (OR, 1.2; 95% CI, 1.1–1.3; P<0.001), self-dislike (OR, 1.2; 95% CI, 1.1–1.3; P<0.001), self-criticalness (OR, 1.2; 95% CI, 1.1–1.3; P<0.001), immobility (OR, 1.5; 95% CI, 1.4–1.6; P<0.001), and anxiety or depression (OR, 1.3; 95% CI, 1.2–1.4; P<0.001) than the concordant–lower severity group (Table S2). For these same questions, symptom minimizers were only more likely to report immobility (OR, 1.1; 95% CI, 1.0–1.2; P=0.004) versus the concordant–lower severity group.

Association of Discordance With Long-Term Clinical Outcomes

The Kaplan-Meier curves for the composite end point and for hospitalization overlapped for participants in the
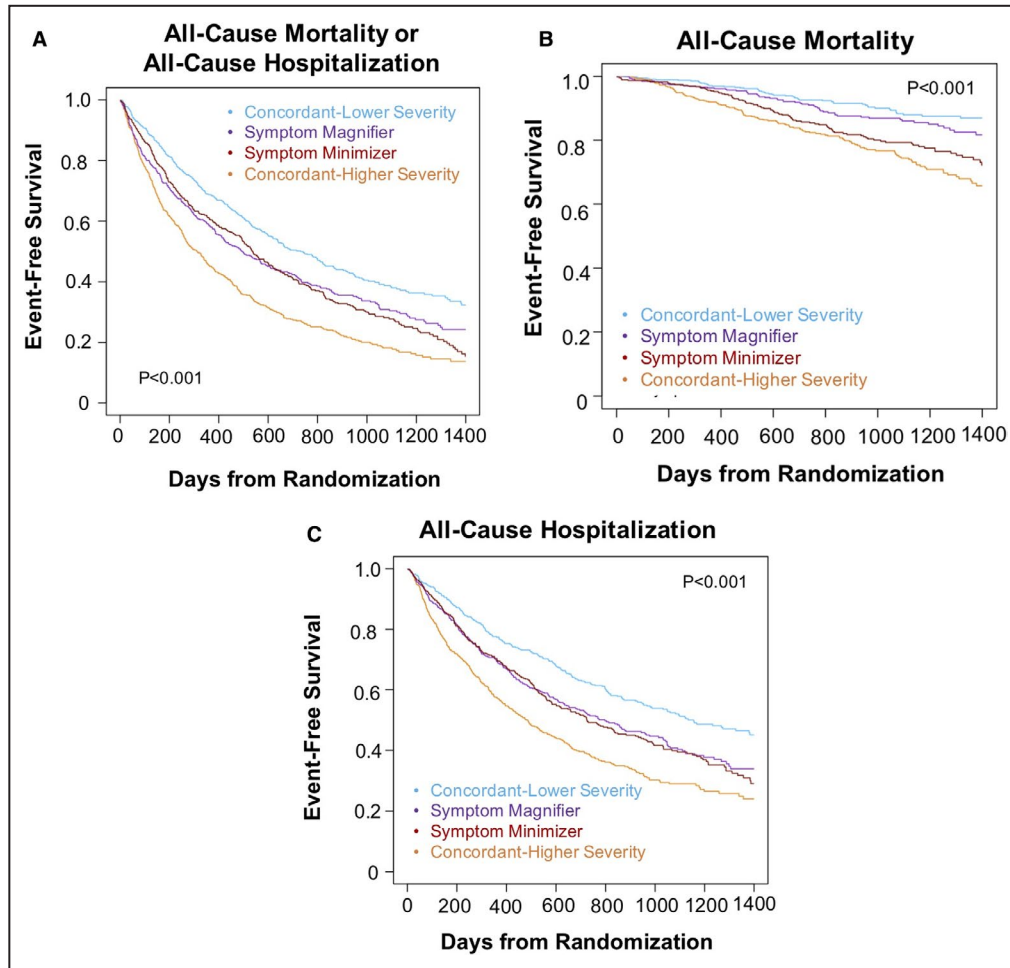
Figure 2. Kaplan-Meier survival curves for time to clinical outcomes stratified by concordant and discordant groups using V_E/VCO_2 slope as the objective measure of heart failure severity.
Kaplan-Meier survival curves are shown for time to (**A**) all-cause mortality or all-cause hospitalization ($P<0.001$ for all comparisons; pairwise $P<0.001$ for symptom magnifiers vs concordant–lower severity group; pairwise $P<0.001$ for symptom minimizers vs concordant–higher severity group), (**B**) all-cause mortality ($P<0.001$ for all comparisons; pairwise $P=0.12$ for symptom magnifiers vs concordant–lower severity group; pairwise $P=0.72$ for symptom minimizers vs concordant–higher severity group), and (**C**) all-cause hospitalization ($P<0.001$ for all comparisons; pairwise $P<0.001$ for symptom magnifiers vs concordant–lower severity group; pairwise $P<0.001$ for symptom minimizers vs concordant–higher severity group). V_E/VCO_2 indicates slope of minute ventilation/carbon dioxide production.

symptom minimizer and magnifier groups (Figure 2A and 2C). Compared with the concordant–lower severity group, symptom magnifiers had a significantly lower probability of survival free of hospitalization and of the composite end point ($P<0.001$ for both comparisons). Relative to the concordant–higher severity groups, symptom minimizers had a significantly higher probability of survival free of hospitalization and of the composite end point ($P<0.001$ for both comparisons). However, there was no difference in all-cause mortality when comparing the symptom magnifiers to concordant–lower severity participants ($P=0.12$) or the symptom minimizers to concordant–higher severity participants ($P=0.072$) (Figure 2B). Using peak VO$_2$ as

the objective measure of heart failure severity, the observations were similar (Figure S2), except the symptom magnifiers had a risk of hospitalization similar to the concordant–lower severity group.

In models adjusted for age, race, and sex, both symptom minimizers and symptom magnifiers, as compared with the concordant–lower severity group, were at higher risk for the composite end point of all-cause mortality or all-cause hospitalization and for hospitalization ($P<0.001$ for both) (Table 3). Symptom minimizers also had a higher risk of all-cause mortality (hazard ratio [HR], 1.79; 95% CI, 1.27–2.50; $P<0.001$). For symptom magnifiers, the higher risk of the composite end point was largely driven by a higher risk of

Table 3.   Cox Proportional Hazards of Long-Term Clinical Outcomes by Concordant and Discordant Groups

	All-Cause Mortality or All-Cause Hospitalization		All-Cause Mortality		All-Cause Hospitalization	
	HR [95% CI]	*P* Value	HR [95% CI]	*P* Value	HR [95% CI]	*P* Value
Concordant–lower severity	Reference	…	Reference	…	Reference	…
Symptom magnifier	1.36 [1.17–1.60]	<0.001	1.39 [0.97–2.00]	0.08	1.43 [1.20–1.71]	<0.001
Symptom minimizer	1.34 [1.14–1.57]	<0.001	1.79 [1.27–2.50]	<0.001	1.40 [1.18–1.68]	<0.001
Concordant–higher severity	1.88 [1.62–2.18]	<0.001	2.41 [1.76–3.31]	<0.001	1.93 [1.64–2.28]	<0.001

Models were adjusted for age, sex, and race. HR indicates hazard ratio.

hospitalization (HR, 1.43; 95% CI, 1.20–1.71; *P*<0.001). Furthermore, despite symptom magnifiers having a KCCQ-CS score 28 points lower (poorer quality of life) than symptom minimizers, their risk of mortality was not increased (HR, 0.79; 95% CI, 0.57–1.1, *P*=0.18, respectively). When we additionally adjusted for depression and obesity, the qualitative results were the same (Table S3).

DISCUSSION

The discordance between patient-reported and objective assessments of severity of illness in ambulatory patients with heart failure is a topic that warrants furthers elucidation. Therefore, in a well-phenotyped cohort, we evaluated this relationship using KCCQ-CS as the patient-reported measure of HR-QOL and V_E/VCO_2 slope as the objective marker of heart failure severity. The major findings of this study are as follows: (1) discordance between patient-reported and objective measures of heart failure severity was common (46%); (2) discordance with symptom magnification relative to objective severity was associated with obesity and a higher burden of adverse psychosocial factors; (3) although hospitalization was associated with patient report, mortality tracked more closely with objective data.

Association of Discordance With Specific Patient Characteristics

Our study significantly extends prior findings on the discordance between HR-QOL and objective measures of heart failure severity[6–11] for several reasons, including that it was based on a large multicenter trial with a well-selected population of medically optimized patients with heart failure with reduced ejection fraction, cardiopulmonary exercise testing studies were interpreted in a core laboratory, and clinical events were adjudicated. Moreover, we identified patient characteristics associated with this discordance, as well as its prognostic implications. These findings are important factors for providers to consider when deciding about

referral and selection of patients with heart failure for cardiac transplantation or left ventricular assist device therapy.[31,32]

The results of our study also highlight that certain functional capacity measures impacted by patient effort should be interpreted within the context of whether discordance exists. Although 6-minute-walk test distance, cardiopulmonary exercise duration, and peak VO_2 have been shown to be independent predictors of survival for patients with heart failure,[33–35] we observed significant differences in these parameters between the respective concordant and discordant groups with the same V_E/VCO_2 slope classification. Therefore, when discordance exists, evaluating objective data that are independent of patient effort, such as the V_E/VCO_2 slope, has particular importance in risk estimation.

Depression and other psychosocial factors seem to be an important driver of the discordance in symptom magnifiers, presumably through influences on patient perception of disease. Depression is known to influence KCCQ[36,37] and has been consistently associated with worse clinical outcomes in patients with heart failure.[38,39] A prior analysis from HF-ACTION demonstrated that depression was associated with lower KCCQ score, but the concept of discordance with respect to an objective measure of heart failure severity was not included in that analysis.[40] This is a critical distinction because some patients with more severe heart failure will have lower HR-QOL related to their more severe illness, and that may result in depression. However, we now demonstrate that depression is associated with lower HR-QOL, even among subjects with objectively less severe heart failure as assessed by lower V_E/VCO_2 slope.

An elevated body mass index and prevalent obesity were also associated with discordance between subjective and objective severity of illness. This finding is not surprising, given that elevated body mass index has been shown to be associated with lower V_E/VCO_2 slope in a prior analysis of HF-ACTION,[41] and because obesity can contribute to the sensation of dyspnea,[42] which may influence self-report of symptoms

on the KCCQ-CS. This observation may have clinical relevance by suggesting that objective assessment of heart failure severity to assess prognosis may be particularly important in obese patients with heart failure.

It is less clear why symptom minimizers seem to underreport symptoms despite having objective markers of more severe heart failure. This may be because symptom minimizers are individuals who have better coping mechanisms or more positive outlooks. However, the psychometric assessments used in HF-ACTION did not evaluate for optimism or a positive demeanor, and thus, we were not able to test that hypothesis.

Discordance and Implications for Clinical Outcomes

Given the present focus on PROs, these observations highlight key insights that clarify the utility of HR-QOL instruments for risk stratification. For example, KCCQ is a validated marker of risk for death or hospitalization in patients with heart failure[43–45]; however, ambulatory patients with heart failure tend to overestimate their life expectancy relative to model-based predictions of survival.[46] Additionally, patients tend to grade their heart failure symptoms differently from their treating physicians.[47] Herein, we demonstrate that in participants with reported HR-QOL discordant to V_E/VCO$_2$ slope, survival tracked more closely with the objective measure of heart failure severity. For example, the symptom minimizers, despite having a KCCQ-CS similar to the concordant–lower severity group, were at a higher risk of mortality. Likewise, despite having a KCCQ-CS 28 points lower, the symptom magnifiers were not at increased risk for mortality as compared with the symptom minimizers.

Hospitalization, in contrast to mortality, was more coupled with the KCCQ-CS. These data are consistent with prior reports showing that low KCCQ predicts hospitalization in patients with heart failure[48,49] but extend those observations by demonstrating that patient report of symptom severity can drive risk of hospitalization even among those with objectively less-severe heart failure. Moreover, our findings suggest that enhanced treatment of depression could lead to a reduction in heart failure hospitalization in some patients. Of note, the symptom minimizers did not have higher mortality as compared with the concordant–higher severity participants, suggesting that their minimization of symptoms did not place them at an unduly higher risk of death.

Limitations

There are limitations to this study. First, there is the possibility of selection bias because patients participated in a clinical trial. Second, because we used the median to divide the sample into concordant and discordant groups based on KCCQ-CS and V_E/VCO$_2$ slope, there is the risk of misclassification, especially in those with scores or values close to the median dichotomous thresholds used for group assignment. However, such misclassification should bias our results to the null. Concomitant pulmonary disease may have contributed to increased symptom perception in the symptom magnifiers. However, V_E/VCO$_2$ slope is a marker of ventilatory efficiency and should reflect such pulmonary influences. Multiple comparisons among the 4 defined groups increase the chance of false-positive associations. However, there was consistency of key findings when tested with an alternative analytic approach (eg, peak VO$_2$ instead of V_E/VCO$_2$ slope). We did not have access to data that could reflect positive psychosocial attributes when characterizing symptom minimizers, and further study is needed to better define that group. The introduction of new therapies for both heart failure and depression since completion of HF-ACTION may limit generalizability of our findings to the contemporary era. Finally, this analysis is based on a static measure of HR-QOL. Further study is needed to determine whether these findings are applicable to serial changes in the KCCQ-CS and VE/VCO$_2$ slope.

CONCLUSIONS

In ambulatory patients with heart failure, discordance between patient-reported HR-QOL and objective measures of disease severity, as assessed by the VE/VCO$_2$ slope, was common. Disproportionately low HR-QOL relative to V_E/VCO$_2$ slope was associated with depression, other adverse psychosocial parameters, and obesity. Further research is needed to determine why some patients minimize their report of symptoms relative to an objective measure of their disease severity. In cases of discordance, objective measures of heart failure severity, and in particular V_E/VCO$_2$ slope, as compared with the KCCQ-CS, were more closely associated with mortality. These findings underscore to practitioners the need to rely not only on patient report but also objective measures of illness severity when risk stratifying their patients with heart failure.

Disclosures

Dr Grodin reports consulting and scientific advisory board fees from Pfizer and Eidos Therapeutics. The remaining authors have no disclosures to report.

Supplementary Material

Tables S1–S3
Figures S1–S2

REFERENCES

1. Ni H, Xu J. Recent trends in heart failure-related mortality: United States, 2000–2014. *NCHS Data Brief*. 2015;231:1–8.

2. Shah KS, Xu H, Matsouaka RA, Bhatt DL, Heidenreich PA, Hernandez AF, Devore AD, Yancy CW, Fonarow GC. Heart failure with preserved, borderline, and reduced ejection fraction: 5-year outcomes. *J Am Coll Cardiol*. 2017;70:2476–2486. DOI: 10.1016/j.jacc.2017.08.074.

3. Steinberg BA, Fang JC. Long-term outcomes of acute heart failure: where are we now? *J Am Coll Cardiol*. 2017;70:2487–2489. DOI: 10.1016/j.jacc.2017.08.075.

4. Burns DJP, Arora J, Okunade O, Beltrame JF, Bernardez-Pereira S, Crespo-Leiro MG, Filippatos GS, Hardman S, Hoes AW, Hutchison S, et al. International Consortium for Health Outcomes Measurement (ICHOM): standardized patient centered outcomes measurement set for heart failure patients. *JACC Heart Fail*. 2020;8:212–222. DOI: 10.1016/j.jchf.2019.09.007.

5. Gluckman TJ, Stevenson LW, Kovacs R. A PROmise to Improve Cardiovascular Care. *J Am Coll Cardiol*. 2019;74:2428–2430. DOI: 10.1016/j.jacc.2019.09.035.

6. Grigioni F, Carigi S, Grandi S, Potena L, Coccolo F, Bacchi-Reggiani L, Magnani G, Tossani E, Musuraca AC, Magelli C, et al. Distance between patients' subjective perceptions and objectively evaluated disease severity in chronic heart failure. *Psychother Psychosom*. 2003;72:166–170. DOI: 10.1159/000069734.

7. Rector TS, Anand IS, Cohn JN. Relationships between clinical assessments and patients' perceptions of the effects of heart failure on their quality of life. *J Card Fail*. 2006;12:87–92. DOI: 10.1016/j.cardfail.2005.10.002.

8. Myers J, Zaheer N, Quaglietti S, Madhavan R, Froelicher V, Heidenreich P. Association of functional and health status measures in heart failure. *J Card Fail*. 2006;12:439–445. DOI: 10.1016/j.cardfail.2006.04.004.

9. Bhardwaj A, Rehman SU, Mohammed AA, Gaggin HK, Barajas L, Barajas J, Moore SA, Sullivan D, Januzzi JL. Quality of life and chronic heart failure therapy guided by natriuretic peptides: results from the ProBNP Outpatient Tailored Chronic Heart Failure Therapy (PROTECT) study. *Am Heart J*. 2012;164:793–799. DOI: 10.1016/j.ahj.2012.08.015.

10. Flynn KE, Lin L, Ellis SJ, Russell SD, Spertus JA, Whellan DJ, Piña IL, Fine LJ, Schulman KA, Weinfurt KP, et al. Outcomes, health policy, and managed care: relationships between patient-reported outcome measures and clinical measures in outpatients with heart failure. *Am Heart J*. 2009;158:S64–S71. DOI: 10.1016/j.ahj.2009.07.010.

11. Lee CS, Hiatt SO, Denfeld QE, Mudd JO, Chien C, Gelow JM. Symptom-hemodynamic mismatch and heart failure event risk. *J Cardiovasc Nurs*. 2015;30:394–402. DOI: 10.1097/JCN.0000000000000175.

12. O'Connor CM, Whellan DJ, Lee KL, Keteyian SJ, Cooper LS, Ellis SJ, Leifer ES, Kraus WE, Kitzman DW, Blumenthal JA, et al. Efficacy and safety of exercise training in patients with chronic heart failure: HF-ACTION randomized controlled trial. *JAMA*. 2009;301:1439–1450. DOI: 10.1001/jama.2009.454.

13. Flynn KE, Piña IL, Whellan DJ, Lin LI, Blumenthal JA, Ellis SJ, Fine LJ, Howlett JG, Keteyian SJ, Kitzman DW, et al. Effects of exercise training on health status in patients with chronic heart failure: HF-ACTION randomized controlled trial. *JAMA*. 2009;301:1451–1459. DOI: 10.1001/jama.2009.457.

14. Whellan DJ, O'Connor CM, Lee KL, Keteyian SJ, Cooper LS, Ellis SJ, Leifer ES, Kraus WE, Kitzman DW, Blumenthal JA, 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. DOI: 10.1016/j.ahj.2006.11.007.

15. Chua TP, Ponikowski P, Harrington D, Anker SD, Webb-Peploe K, Clark AL, Poole-Wilson PA, Coats AJS. Clinical correlates and prognostic significance of the ventilatory response to exercise in chronic heart failure. *J Am Coll Cardiol*. 1999;29:1585–1590. DOI: 10.1016/S0735-1097(97)00078-8.

16. Mehra MR, Canter CE, Hannan MM, Semigran MJ, Uber PA, Baran DA, Danziger-Isakov L, Kirklin JK, Kirk R, Kushwaha SS, et al. The 2016 International Society for Heart Lung Transplantation listing criteria for heart transplantation: a 10-year update. *J Heart Lung Transplant*. 2016;35:1–23. DOI: 10.1016/j.healun.2015.10.023.

17. Arena R, Myers J, Guazzi M. The clinical and research applications of aerobic capacity and ventilatory efficiency in heart failure: an evidence-based review. *Heart Fail Rev*. 2008;13:245–269. DOI: 10.1007/s10741-007-9067-5.

18. Sun X-G, Hansen JE, Beshai JF, Wasserman K. Oscillatory breathing and exercise gas exchange abnormalities prognosticate early mortality and morbidity in heart failure. *J Am Coll Cardiol*. 2010;55:1814–1823. DOI: 10.1016/j.jacc.2009.10.075.

19. Ferreira AM, Tabet J-Y, Frankenstein L, Metra M, Mendes M, Zugck C, Beauvais F, Cohen-Solal A. Ventilatory efficiency and the selection of patients for heart transplantation. *Circ Heart Fail*. 2010;3:378–386. DOI: 10.1161/CIRCHEARTFAILURE.108.847392.

20. Malhotra R, Bakken K, D'Elia E, Lewis GD. Cardiopulmonary exercise testing in heart failure. *JACC Heart Fail*. 2016;4:607–616. DOI: 10.1016/j.jchf.2016.03.022.

21. Agostoni P, Corrà U, Cattadori G, Veglia F, Battaia E, La Gioia R, Scardovi AB, Emdin M, Metra M, Sinagra G, et al. Prognostic value of indeterminable anaerobic threshold in heart failure. *JACC Heart Fail*. 2013;6:977–987. DOI: 10.1161/CIRCHEARTFAILURE.113.000471.

22. Balady GJ, Arena R, Sietsema K, Myers J, Coke L, Fletcher GF, Forman D, Franklin B, Guazzi M, Gulati M, et al. Clinician's guide to cardiopulmonary exercise testing in adults: a scientific statement from the American Heart Association. *Circulation*. 2010;122:191–225. DOI: 10.1161/CIR.0b013e3181e52e69.

23. Arena R, Myers J, Guazzi M. Cardiopulmonary exercise testing is a core assessment for patients with heart failure. *Congest Heart Fail*. 2011;17:115–119. DOI: 10.1111/j.1751-7133.2011.00216.x.

24. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. *J Am Coll Cardiol*. 2000;35:1245–1255. DOI: 10.1016/S0735-1097(00)00531-3.

25. Chua TP, Ponikowski P, Harrington D, Anker SD, Webb-Peploe K, Clark AL, Poole-Wilson PA, Coats AJS. Clinical correlates and prognostic significance of the ventilatory response to exercise in chronic heart failure. *J Am Coll Cardiol*. 1997;29:1585–1590. DOI: 10.1016/S0735-1097(97)00078-8.

26. Mancini DM, Eisen H, Kussmaul W, Mull R Jr, Edmunds LH, Wilson JR. Value of peak exercise oxygen consumption for optimal timing of cardiac transplantation in ambulatory patients with heart failure. *Circulation*. 1991;83:778–786. DOI: 10.1161/01.CIR.83.3.778.

27. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clin Psychol Rev*. 1988;8:77–100. DOI: 10.1016/0272-7358(88)90050-5.

28. Zimet G, Powell S, Farley G, Werkman S, Berkoff K. Psychometric characteristics of the Multidimensional Scale of Perceived Social Support. *J Pers Assess*. 1990;55:610–617. DOI: 10.1207/s15327752jpa5503&4_17.

29. Stamford BA. Validity and reliability of subjective ratings of perceived exertion during work. *Ergonomics*. 1976;19:53–60. DOI: 10.1080/00140137608931513.

30. Calvert MJ, Freemantle N, Cleland JGF. The impact of chronic heart failure on health-related quality of life data acquired in the baseline phase of the CARE-HF study. *Eur J Heart Fail*. 2005;7:243–251. DOI: 10.1016/j.ejheart.2005.01.012.

31. Allen LA, Gheorghiade M, Reid KJ, Dunlay SM, Chan PS, Hauptman PJ, Zannad F, Konstam MA, Spertus JA. Identifying patients hospitalized with heart failure at risk for unfavorable future quality of life. *Circ Cardiovasc Qual Outcomes*. 2011;4:389–398. DOI: 10.1161/CIRCOUTCOMES.110.958009.

32. Faranoff AC, DeVore AD, Mentz RJ, Daneshmand MA, Patel CB. Patient selection for advanced heart failure therapy referral. *Crit Pathw Cardiol*. 2014;13:1–5. DOI: 10.1097/HPC.0000000000000004.

33. Forman DE, Fleg JL, Kitzman DW, Brawner CA, Swank AM, McKelvie RS, Clare RM, Ellis SJ, Dunlap ME, Bittner V. 6-min walk test provides prognostic utility comparable to cardiopulmonary exercise testing in ambulatory outpatients with systolic heart failure. *J Am Coll Cardiol*. 2012;60:2653–2661. DOI: 10.1016/j.jacc.2012.08.1010.

34. Keteyian SJ, Patel M, Kraus WE, Brawner CA, McConnell TR, Piña IL, Leifer ES, Fleg JL, Blackburn G, Fonarow GC, et al. Variables measured during cardiopulmonary exercise testing as predictors of mortality in chronic heart failure. *J Am Coll Cardiol*. 2016;67:780–789. DOI: 10.1016/j.jacc.2015.11.050.

35. Gitt AK, Wasserman K, Kilkowski C, Kleemann T, Kilkowski A, Bangert M, Schneider S, Schwarz A, Senges J. Exercise anaerobic threhold and ventilatory efficiency identify heart failure patients for high risk of early death. *Circulation*. 2002;106:3079–3084. DOI: 10.1161/01.cir.0000041428.99427.06.

36. Sullivan M, Levy WC, Russo JE, Spertus JA. Depression and health status in patients with advanced heart failure: a prospective study in tertiary care. *J Card Fail*. 2004;10:390–396. DOI: 10.1016/j.cardfail.2004.01.011.

37. Faller H, Störk S, Schuler M, Schowalter M, Steinbüchel T, Ertl G, Angermann CE. Depression and disease severity as predictors of health-related quality of life in patients with chronic heart failure—a structural equation modeling approach. *J Card Fail*. 2009;15:286–292. DOI: 10.1016/j.cardfail.2008.10.022.

38. Rutledge T, Reis VA, Linke SE, Greenberg BH, Mills PJ. Depression in heart failure a meta-analytic review of prevalence, intervention effects, and associations with clinical outcomes. *J Am Coll Cardiol*. 2006;48:1527–1537. DOI: 10.1016/j.jacc.2006.06.055.

39. Moraska AR, Chamberlain AM, Shah ND, Vickers KS, Rummans TA, Dunlay SM, Spertus JA, Weston SA, McNallan SM, Redfield MM, et al. Depression, healthcare utilization, and death in heart failure: a community study. *Circ Heart Fail*. 2013;6:387–394. DOI: 10.1161/CIRCHEARTFAILURE.112.000118.

40. Gottlieb SS, Kop WJ, Ellis SJ, Binkley P, Howlett J, O'Connor C, Blumenthal JA, Fletcher G, Swank AM, Cooper L, et al. Relation of depression to severity of illness in heart failure (from Heart Failure And a Controlled Trial Investigating Outcomes of Exercise Training [HF-ACTION]). *Am J Cardiol*. 2009;103:1285–1289. DOI: 10.1016/j.amjcard.2009.01.025.

41. Horwich TB, Leifer ES, Brawner CA, Fitz-Gerald MB, Fonarow GC. The relationship between body mass index and cardiopulmonary exercise testing in chronic systolic heart failure. *Am Heart J*. 2009;158:S31–S36. DOI: 10.1016/j.ahj.2009.07.016.

42. Bernhardt V, Babb TG. Exertional dyspnoea in obesity. *Eur Respir Rev*. 2016;25:487–495. DOI: 10.1183/16000617.0081-2016.

43. Soto GE, Jones P, Weintraub WS, Krumholz HM, Spertus JA. Prognostic value of health status in patients with heart failure after acute myocardial infarction. *Circulation*. 2004;110:546–551. DOI: 10.1161/01.CIR.0000136991.85540.A9.

44. Kosiborod M, Soto GE, Jones PG, Krumholz HM, Weintraub WS, Deedwania P, Spertus JA. Identifying heart failure patients at high risk for near-term cardiovascular events with serial health status assessments. *Circulation*. 2007;115:1975–1981. DOI: 10.1161/CIRCULATIONAHA.106.670901.

45. Pokharel Y, Khariton Y, Tang Y, Nassif ME, Chan PS, Arnold SV, Jones PG, Spertus JA. Association of serial Kansas City Cardiomyopathy Questionnaire assessments with death and hospitalization in patients with heart failure with preserved and reduced ejection fraction: a secondary analysis of 2 randomized clinical trials. *JAMA Cardiol*. 2017;2:1315–1321. DOI: 10.1001/jamacardio.2017.3983.

46. Allen LA, Yager JE, Funk MJ, Levy WC, Tulsky JA, Bowers MT, Dodson GC, O'Connor CM, Felker GM. Discordance between patient-predicted and model-predicted life expectancy among ambulatory patients with heart failure. *JAMA*. 2008;299:2533–2542. DOI: 10.1001/jama.299.21.2533.

47. Goode KM, Nabb S, Cleland JGF, Clark AL. A comparison of patient and physician-rated New York Heart Association class in a community-based heart failure clinic. *J Card Fail*. 2008;14:379–387. DOI: 10.1016/j.cardfail.2008.01.014.

48. Heidenreich PA, Spertus JA, Jones PG, Weintraub WS, Rumsfeld JS, Rathore SS, Peterson ED, Masoudi FA, Krumholz HM, Havranek EP, et al. Health status identifies heart failure outpatients at risk for hospitalization or death. *J Am Coll Cardiol*. 2005;47:752–756. DOI: 10.1016/j.jacc.2005.11.021.

49. Dunlay SM, Gheorghiade M, Reid KJ, Allen LA, Chan PS, Hauptman PJ, Zannad F, Maggioni AP, Swedberg K, Konstam MA, et al. Critical elements of clinical follow-up after hospital discharge for heart failure: insights from the EVEREST trial. *Eur J Heart Fail*. 2010;12:367–374. DOI: 10.1093/eurjhf/hfq019.

SUPPLEMENTAL MATERIAL

Table S1. Association of Concordant and Discordant Groups with Perceived Limitations.

	Beck Depression Inventory II Score		Multidimensional Scale of Perceived Social Support		Borg Rating at Peak Exercise	
	ß Coefficient [95% CI]	P	ß Coefficient [95% CI]	P	ß Coefficient [95% CI]	P
Concordant-Lower Severity	Ref.	---	Ref.	---	Ref.	---
Symptom Magnifier	7.34 [6.43, 8.26]	<0.001	-0.46 [-0.61, -0.31]	<0.001	0.38 [0.08, 0.67]	0.012
Symptom Minimizer	0.19 [-0.76, 1.13]	0.70	0.00 [-0.15, 0.15]	0.98	0.05 [-0.25, 0.36]	0.74
Concordant-Higher Severity	7.54 [6.64, 8.44]	<0.001	-0.31 [-0.46, -0.17]	<0.001	0.54 [0.25, 0.82]	<0.001

95% CI are displayed as [lower bound, upper bound].
Models were adjusted for age, sex, and race.

Table S2. Association of Concordant and Discordant Groups with Baseline Perception of Limitations as Assessed by BDI II and EuroQoL Questions.

| | Beck Depression Inventory II | | | | | | | | EuroQoL | | | |
| | Pessimism | | Past Failure | | Self-Dislike | | Self-Criticalness | | Immobility | | Anxiety / Depression | |
	OR [95% CI]	P	OR [95% CI]	P	OR [95% CI]	P	OR [95% CI]	P	OR [95% CI]	P	OR [95% CI]	P
Concordant-Lower Severity	Ref.	---	Ref.	---	Ref.	---	Ref.	---	Ref.	---	Ref.	---
Symptom Magnifier	1.3 [1.2-1.3]	<0.001	1.2 [1.1-1.3]	<0.001	1.2 [1.1-1.3]	<0.001	1.2 [1.1-1.3]	<0.001	1.5 [1.4-1.6]	<0.001	1.3 [1.2-1.4]	<0.001
Symptom Minimizer	1.0 [0.9-1.0]	0.06	1.0 [0.9-1.0]	0.42	1.0 [0.9-1.1]	0.87	1.0 [0.9-1.1]	0.83	1.1 [1.0-1.2]	0.004	1.0 [0.9-1.0]	0.53
Concordant-Higher Severity	1.3 [1.2-1.3]	<0.001	1.1 [1.1-1.2]	<0.001	1.2 [1.1-1.3]	<0.001	1.2 [1.1-1.3]	<0.001	1.6 [1.5-1.7]	<0.001	1.3 [1.2-1.4]	<0.001

Models were adjusted for age, sex, and race.

Table S3. Cox Proportional Hazards of Long-term Clinical Outcomes by Concordant and Discordant Groups.

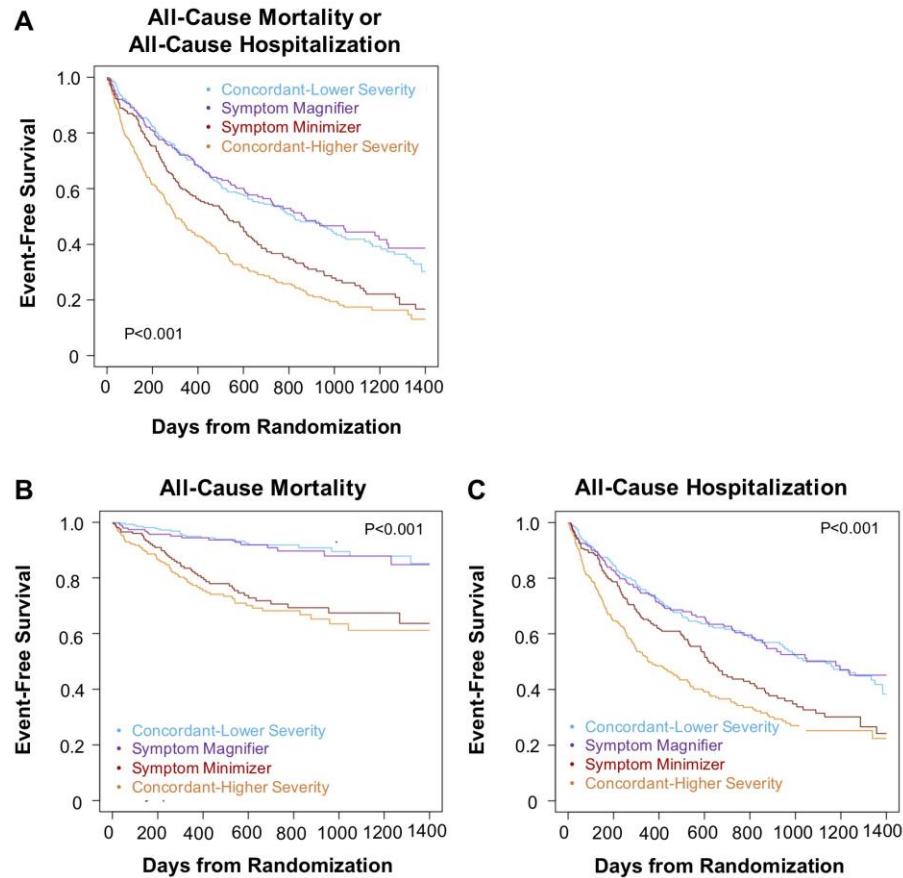|	All-Cause Mortality or All-Cause Hospitalization		All-Cause Mortality		All-Cause Hospitalization	
	HR [95% CI]	P	HR [95% CI]	P	HR [95% CI]	P
Concordant-Lower Severity	Ref.	---	Ref.	---	Ref.	---
Symptom Magnifier	1.31 [1.12-1.53]	<0.001	1.42 [0.98-2.04]	0.063	1.39 [1.17-1.66]	<0.001
Symptom Minimizer	1.37 [1.17-1.60]	<0.001	1.79 [1.28-2.51]	<0.001	1.42 [1.19-1.70]	<0.001
Concordant-Higher Severity	1.86 [1.60-2.15]	<0.001	2.43 [1.77-3.34]	<0.001	1.92 [1.63-2.26]	<0.001

Models were adjusted for age, sex, race, depression, and obesity.

Figure S1. Scatter Plot Examining the Relationship between KCCQ Clinical Summary Score and V_E/VCO_2 Slope.
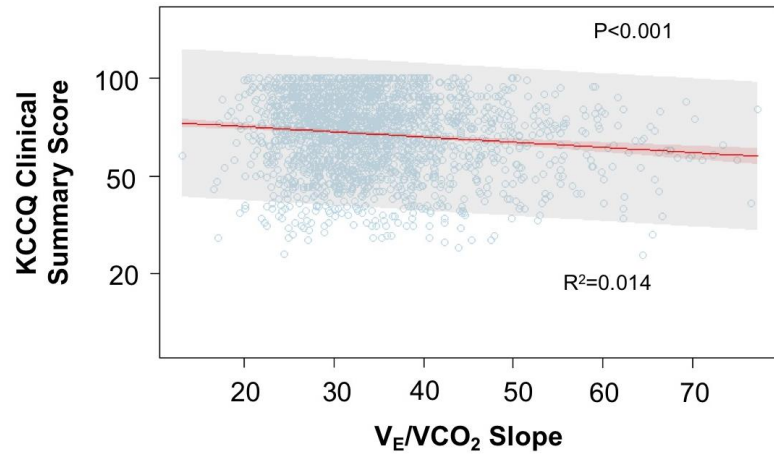
The fitted line represents the linear regression model. The overall model was statistically significant (P<0.001) but explained only

1.4% of this relationship (R^2=0.014)

Figure S2. Kaplan-Meier Survival Curves for Time to Clinical Outcomes Stratified by Concordant and Discordant Groups Using Peak VO₂ as the Objective Measure of Heart Failure Severity.

Kaplan-Meier survival curves are shown for time to A) all-cause mortality or all-cause hospitalization (P <0.001 for all comparisons), B) all-cause mortality (P <0.001 for all comparisons), and C) all-cause hospitalization (P<0.001 for all comparisons)