ORIGINAL RESEARCH

Interpretability of the European Heart Failure Self-care Behaviour scale

Kim P Wagenaar[1]
Berna DL Broekhuizen[1]
Frans H Rutten[1]
Anna Strömberg[2]
Henk F van Stel[1]
Arno W Hoes[1]
Tiny Jaarsma[2]

[1]Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, the Netherlands; [2]Department of Social and Welfare Studies, Linköping University, Linköping, Sweden

Objective: The European Heart Failure Self-care Behaviour scale (EHFScBs) is a valid patient-reported questionnaire to measure self-care behavior of heart failure (HF) patients. We assessed the interpretability of the EHFScBs.

Methods: We used data of 1,023 HF patients. Interpretability refers to the clinical meaning of the score and its changes over time. We operationalized interpretability by evaluating distributions of EHFScBs scores across relevant HF subgroups by eyeballing, by testing the risk on hospitalizations and mortality of a plausible threshold, and by determining a clinically relevant minimal important change (MIC). The scale score ranged from 0 to 100, with a higher score meaning better self-care. A threshold of ≥70 was defined as adequate and <70 as inadequate self-care.

Results: The EHFScBs scores were similarly normally distributed among the subgroups with a mean between 57.8 (SD 19.4) and 72.0 (SD 18.0). The 464 HF patients with adequate self-care had significantly less all-cause hospitalizations than the 559 patients with inadequate self-care.

Conclusion: The degree of self-care showed to be independent of relevant HF subgroups. A single threshold of 70 accurately discriminated between patients with adequate and inadequate self-care.

Practice implications: The threshold of 70 can be used in designing studies and informing health policy makers.

Keywords: heart failure, self-care, interpretability, patient-reported outcome, threshold and minimal important change

Introduction

The European Heart Failure Self-care Behaviour Scale (EHFScBs) is a patient-reported outcome (PRO) questionnaire to measure self-care behavior of heart failure (HF) patients. Self-care behavior reflects actions undertaken to maintain life, healthy functioning, and well-being.[1,2]

The EHFScBs has been psychometrically tested,[1–3] ie, its validity (does it measure what it should measure?), reliability (is it free of measurement error?), and responsiveness (is it sensitive to changes?) have been evaluated.[4–6] Besides these test characteristics, the interpretability is an important and recently acknowledged PRO characteristic that – as for many other PROs – has not yet been evaluated for the EHFScBs.[4,5] Interpretability refers to the possible clinical meaning of the score results, and its changes over time.[4,5] Despite the fact that the EHFScBs is widely used as an outcome in clinical trials, the clinical relevance of the score results and the changes in the score over time are unknown. Moreover, a threshold for "adequate self-care" related to relevant outcomes is unknown.

Correspondence: Kim P Wagenaar
Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, Universiteitsweg 100, 3584 CG Utrecht, the Netherlands
Tel +31 8 8756 9636
Email k.wagenaar-2@umcutrecht.nl

Knowledge of the interpretability of the EHFScBs score is urgently needed to properly value results of clinical trials in HF using this PRO as an outcome and to address whether the scale is suitable for evaluation of health care interventions.[7]

Our overall study aim was to assess the interpretability of the EHFScBs in a cohort of patients with HF. The study questions were as follows:

1. What is the distribution of the scores of the EHFScBs in HF patients?
2. Can a threshold of the EHFScBs scores distinguish between adequate and inadequate self-care in HF patients?
3. What is the clinically relevant minimal important change (MIC) in EHFScBs score in HF patients?

Methods

Cross-sectional and longitudinal data of participants of the "Coordinating study evaluating Outcomes of Advising and Counselling in Heart Failure" (COACH) study[8] were used. Missing values of variables were imputed by multiple imputation using SPSS version 21.0 (IBM Corporation, Armonk, NY, USA).[9]

Study population

The COACH study was a three-arm multicenter, randomized clinical trial in patients with HF, comparing the effects of moderate and intensive disease management support to usual care. Details of the COACH study have been published elsewhere.[8,10] Patients, 18 years or older, were included during hospitalization for HF and were symptomatic (based on New York Heart Association functional classification [NYHA], classes II–IV). Important exclusion criteria were: 1) inability to complete the questionnaires; 2) invasive procedures (eg, cardiac surgery or intervention) within the last 6 months or planned within the next 3 months; and 3) ongoing evaluation for heart transplantation. During the hospitalization, patients were randomly allocated to one of the three arms.

All patients received "routine" care by their cardiologist common in the Netherlands. This included a visit at the outpatient clinic within 2 months after hospital discharge and every 6 months after that. Patients in the two intervention arms with extra support were visited by an HF nurse during hospitalization and received education and extra support. They were also scheduled for additional visits at the outpatient clinic. Patients in the most intensive support arm had, on top of the latter, weekly contact by telephone with the HF nurse in the first month after hospital discharge and monthly contact afterward including two home visits.[8,10]

Patients were followed for 18 months. At baseline, patient characteristics (eg, age, educational level, depressive symptoms) were measured. At baseline and after 12 months, all participants filled out questionnaires on self-care behavior (EHFScBs) and HF-related quality of life (Minnesota Living with Heart Failure Questionnaire [MLHFQ]). The NYHA classification was collected from the hospital medical file.[8] In addition, at these time points, B-type natriuretic peptide (BNP) levels were measured, and hospitalizations and death were retrieved from the hospital medical files at the end of the study and adjudicated by a blinded end point committee.

The medical ethics committee of the University Medical Center Groningen approved the study protocol, and all patients provided written informed consent.

EHFScBs

The EHFScBs includes both self-reported consulting (eg, "if I gain weight I contact a doctor or nurse") and adherence to regimen behaviors (eg, "I exercise regularly"). It consists of nine items that are scored on a 5-point Likert scale, resulting in a standardized score from 0 to 100 (every item is given an equal weight) with a higher score meaning better self-care.[1,3] The EHFScBs was developed in 2003[2] and reduced to a nine-item scale (EHFScB-9) in 2009.[1]

Operationalizing interpretability

We operationalized interpretability by evaluating distributions of EHFScBs scores across relevant subgroups, by evaluating whether a plausible threshold in the EHFScBs scores can distinguish between adequate and inadequate self-care, and by assessing a clinically relevant MIC.[5,11–14]

Information on how the EHFScBs scores of the study population and its relevant subgroups are distributed is important for the interpretation of the scores. It provides information on whether the population and its subgroups have a low or a high score on average and if the scores are distributed over the whole range of the scale or clustered. We considered the following subgroups as clinically relevant and possibly related to self-care: age (<60 years, between 60 and 80 years, ≥80 years), gender (female, male), marital status (having a partner, single), level of education (primary, secondary and vocational, middle and higher education), disease severity measured with the NYHA classification (I, II, III, IV), duration of HF (0–3 months, 3–6 months, 6 months–1 year, 1–2 years, 2–3 years, 3–5 years, >5 years), left ventricular ejection fraction (LVEF; <20, between 20 and 35, ≥35),

concurrent depressive symptoms (not depressed, depressed), and patient's knowledge on HF (insufficient, sufficient).[5]

To determine whether a threshold in scores distinguishes between adequate and inadequate self-care, we defined this as 70, based on the threshold of another self-care scale, the "Self-Care of Heart Failure Index" (SCHFI). For the SCHFI, a threshold for adequate self-care was set at ≥70 based on narrative accounts.[15] The SCHFI and the EHFScBs measure similar constructs/aspects and have the same standardized score (from 0 to 100).[3]

To assess the MIC, we used the change in EHFScBs score between baseline and 12 months. For this analysis, we included patients with an EHFScBs score of <70 at baseline, since these patients still have room for improvement in self-care. To help "quantify" the change in score over time, we used an anchor-based method. In this approach, one or more anchors (external criterion, a kind of reference standard) are used, which is an easy to interpret variable that is clearly related to prognosis, either improvement or deterioration. In this case, the anchor-based method assesses which changes in self-care correspond with the clinically relevant changes in the anchor variable.[5] Anchors used were health status, HF-related hospitalizations, and BNP as the biomarker level of left ventricular wall stress.[16]

Details on the anchors

We used the MLHFQ to measure HF-related health status. It is a validated questionnaire including 21 items on patient perceptions on the effects of HF on their physical, psychological, and socioeconomic functioning on a 6-point scale with higher scores indicating worse quality of life.[17] Based on the A-HeFT study, the following clinically relevant changes were determined: improved (≥−5 units), unchanged (−4 up to 4 units), and worse (≥5 units).[18] HF-related hospitalization was defined as unplanned overnight stay in a hospital due to progression of HF or directly related to HF.[8] Categories used for HF-related hospitalization were no hospitalization and one, two, three, four, five, and six or more hospitalizations. Finally, we used the level and change of BNP. We hypothesized that adequate self-care results in maintenance of volume homeostasis by appropriate (diuretic) drug use and, thus, lower left ventricular wall stress and BNP values.[16,19] An increase and decrease in BNP level of 25, 25–50, or >50% were considered as clinically relevant.[20]

Analysis of the EHFScBs score

For the distributions of the EHFScBs score, we used data at 12 months follow-up. These data, and not baseline data,

were used to avoid homogeneous results in self-care due to recent hospitalization.

At 12 months follow-up, most participants were at home and consequently not reminded on a daily basis by the nurses to perform self-care behaviors like taking medication. Using data collected at 12 months provides probably a more realistic view on self-care than using data collected at baseline. The distributions are presented by histograms.

We tested the discriminative value of the threshold 70 by dividing the study population into two groups at baseline; those with a score <70 and ≥70 on the EHFScBs. Participants considered as having adequate self-care (EHFScBs score ≥70) were compared with those with inadequate self-care (EHFScBs score <70) regarding all-cause mortality and all- and HF-related hospitalizations at 12 months, by a chi-square test.

To determine the MIC, first the change in self-care as measured with EHFScBs was correlated with the change over 1 year in the three anchors. If the correlation was sufficient (at least $r=0.3/-0.3$[21]), the MIC was calculated for the separate anchors. To calculate the MIC, we applied the most frequently used method,[14] which is called the "mean change method". In the mean change method, the MIC is calculated from the mean change in the score on the EHFScBs in patients whose values are "minimally importantly changed" between baseline and 12 months follow-up according to the anchor.

Results

The mean age of the 1,023 HF patients was 70.8 years (SD 11.4 years), and 63% were men. The mean LVEF was 33.8% (SD 13.7), and 96% of the patients were classified as NYHA II or III at discharge from the hospital. Most (61%) were living with a partner and reported a low education (86%; Table 1).

Distribution and threshold

The scores on the EHFScBs for the 1,023 HF patients were normally distributed with generally higher scores (mean 65.3 [SD 20.3]), meaning better self-care. The EHFScBs scores among most subgroups (gender, NHYA class, LVEF, duration of HF, depressive symptoms, educational level and marital status [Figure 1]) were distributed similarly. The mean score per subgroup ranged from 57.8 (SD 19.4) to 72.0 (SD 18.0). However, small differences in age groups were present with a mean score in self-care of 72 (SD 18.0), 66 (SD 20.2), and 59 (SD 20.1) for <60 years, between 60 and 80 years, and ≥80 years of age, respectively. Differences

Table 1 Baseline characteristics of the 1,023 COACH study participants

	n	%
Mean age in years (SD)	1,023	70.8 (11.4)
Male gender	639	62.5
Mean LVEF% (SD)	1,023	33.8 (13.7)
NYHA		
2	514	50.2
3	471	46.0
4	38	3.7
Length of HF in months	1,023	Median 3.6 (IQR 0.7–46.5)
Depressive symptoms, CES-D score ≥16	405	39.6
Marital status		
Living with a partner	619	60.5
Single	404	39.5
Educational level		
Primary education	359	35.1
Secondary education and lower and secondary vocational education	516	50.4
Middle and higher education	148	14.5
Mean knowledge in HF (SD)	1,023	11.0 (2.3)
Management randomized to		
Care as usual	338	30.0
Basic support	341	33.3
Intensive support	344	33.6

Abbreviations: COACH, Coordinating study evaluating Outcomes of Advising and Counseling in Heart Failure; IQR, interquartile range; LVEF, left ventricular ejection fraction; NYHA, New York Heart Association functional classification; HF, heart failure; CES-D, Center for Epidemiological Studies Depression scale.

were also present in high and low HF knowledge groups with a mean of 67 (SD 20.0) and 60 (SD 20.2), respectively.

Regarding the threshold, HF patients with adequate self-care had significantly less hospitalizations due to all-cause as compared to the patients with inadequate self-care; OR 0.77 (95% CI 0.60–0.99; Table 2).

MIC

Because the correlation coefficients for the change in the EHFScBs and the concomitant change in the anchors did not exceed 0.3 (poor correlation), the MIC could not be calculated. Figure 2 shows the associations by scatter plots with the concomitant correlation coefficient (*r*).

Discussion and conclusion

The scores on the EHFScBs for the study sample were overall normally distributed with similar score distributions among subgroups. The mean score per subgroup ranged from 57.8 (SD 19.4) to 72.0 (SD 18.0). HF patients with adequate self-care (EHFScBs score ≥70) had significantly less all-cause hospitalizations than patients considered to apply inadequate self-care (EHFScBs score <70), OR 0.77 (95% CI 0.60–0.99). The correlation of the change in the EHFScBs over 12 months with the change in the results of the used anchors was poor and did not allow computing the clinical MIC for the EHFScBs score.

The distribution of the EHFScBs scores in the most relevant subgroups was similar to the distribution of the whole study population. This means that despite gender, depressive symptoms, LVEF, NYHA class, education, marital status, or duration of HF self-care is similarly distributed ranging from 11 to 100 with a mean score of 65.3 (SD 20.3). However, if patients are older or if their HF knowledge is insufficient, their self-care is somewhat lower. The latter findings are in line with the literature.[22]

Our study showed that the degree of self-care seems to be independent of relevant HF subgroup characteristics. However, one could argue that these findings might be less homogeneous in other study samples. It is common that trial participants as in the COACH trial (randomized controlled trial [RCT]) are a selective sample and mainly represent those motivated to improve the subjects under study, in this case self-care.[23,24] In addition, our study population had been hospitalized and educated previously, which most probably also contributed to a homogeneous population regarding self-care abilities.

We were surprised to find no correlation of the EHFScBs with any of the anchors. Especially with HF-related hospitalizations, we would have expected to see a correlation as a previous review found that disease management programs enhancing patient self-care activities reduced HF-related hospitalizations significantly.[25] However, the studies included in this review assessed self-care with another self-care scale or a previous version of the EHFScBs, which may explain the difference in correlation.[25] As a consequence of this lack of correlation, we could not calculate the MIC. An MIC would have provided useful information. For example, if we had found that self-care decreased by 2 points (on average) following every additional hospitalization, we could have concluded that because hospitalization is a major event, a decrease in 2 points on the EHFScBs is also a clinically relevant major decrease.

For future research, we recommend to use anchors that are by content much closer to self-care. One could think of a measure of "adherence to HF treatment" or "HF consulting behaviors", as these concepts are components of self-care. However, even though these measures are probably much stronger related to self-care, the interpretation might be more challenging (while straightforward interpretation is a prerequisite for an anchor) than clinical measures. Another solution may be to create a composite anchor, containing
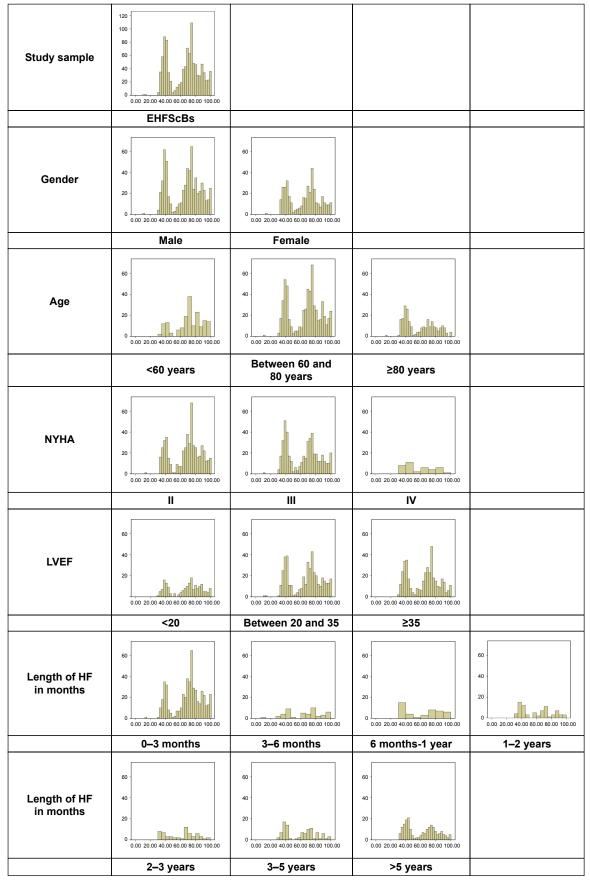
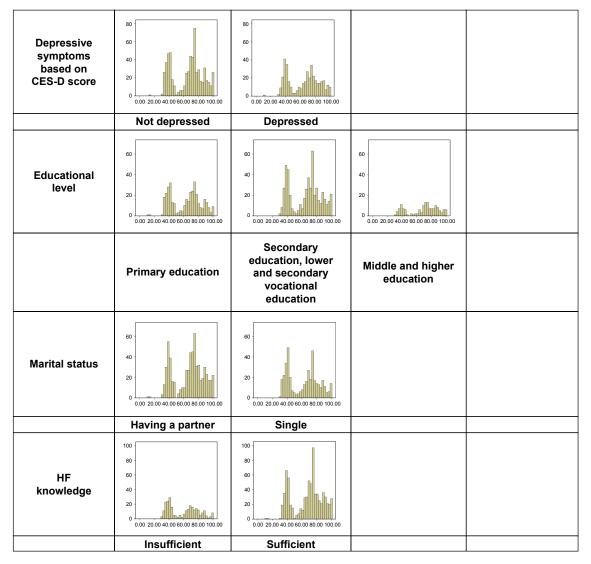|			
Study sample			
	EHFScBs		
Gender			
	Male	**Female**	
Age			
	<60 years	**Between 60 and 80 years**	**≥80 years**
NYHA			
	II	**III**	**IV**
LVEF			
	<20	**Between 20 and 35**	**≥35**

Length of HF in months				
	0–3 months	**3–6 months**	**6 months-1 year**	**1–2 years**

Length of HF in months			
	2–3 years	**3–5 years**	**>5 years**

Figure 1 (*Continued*)

Figure 1 Distribution of the EHFScBs scores at 12 months in relevant subgroups of HF (*y*-axis, number of patients; *x*-axis, EHFScBs score [0–100]).
Abbreviations: EHFScBs, European Heart Failure Self-care Behaviour scale; HF, heart failure; NYHA, New York Heart Association functional classification; LVEF, left ventricular ejection fraction; CES-D, Centre for Epidemiologic Studies Depression scale.

multiple anchors, for example including the ones we used. Such composite anchors as a reference standard for self-care do not yet exist. However, composite outcomes have been used in previous HF research.[26]

Strengths and limitations

Before analysis, missing values were imputed with multiple imputation with SPSS. This is a strength of our study as analyses can be performed on the entire dataset and not only

Table 2 Relation between adequate (EHFScBs score ≥70) versus inadequate self-care (EHFScBs score <70) and all-cause mortality and hospitalizations during 12 months of follow-up in 1,023 patients with HF

Outcomes	Adequate self-care (n=464)	Inadequate self-care (n=559)	*P*-value[a]
	n (%)	n (%)	
All-cause mortality	117 (25.2)	155 (27.7)	0.37
All-cause hospitalizations	241 (51.9)	326 (58.3)	**0.04***
HF-related hospitalizations	113 (24.4)	147 (26.3)	0.48

Notes: [a]*P*-value was calculated with the Pearson's chi-square test. *Signficant *P*-value.
Abbreviations: EHFScBs, European Heart Failure Self-care Behaviour scale; HF, heart failure.

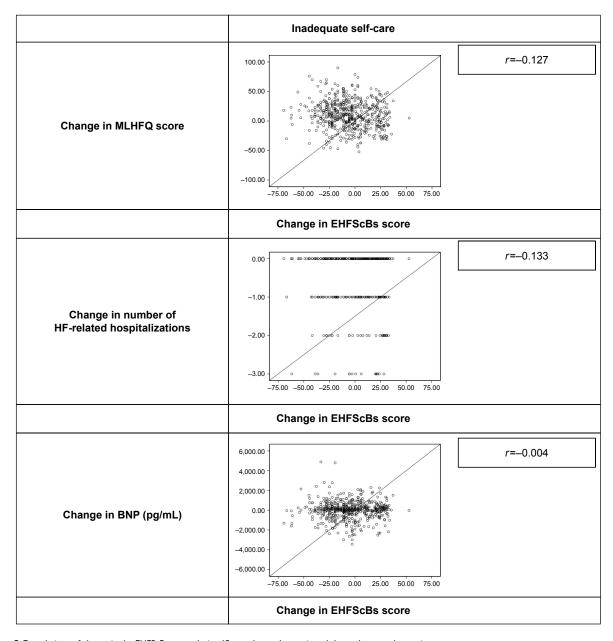|	**Inadequate self-care**
Change in MLHFQ score	 *r*=−0.127 **Change in EHFScBs score**
Change in number of HF-related hospitalizations	 *r*=−0.133 **Change in EHFScBs score**
Change in BNP (pg/mL)	 *r*=−0.004 **Change in EHFScBs score**

Figure 2 Correlations of change in the EHFScBs score during 12 months on the *x*-axis and the anchors on the *y*-axis.
Abbreviations: EHFScBs, European Heart Failure Self-care Behaviour scale; MLHFQ, Minnesota Living with Heart Failure Questionnaire; HF, heart failure; BNP, B-type natriuretic peptide.

on the participants with complete data. In addition, results are more trustworthy than with complete case analysis.[27,28]

Unfortunately, in our data, we did not have information on other appropriate anchors, for example, measures closer to the concept of self-care, like compliance to HF treatment regimens.[1,29] Therefore, we could not incorporate these measures in our analysis.

Practice implications

The variance in the distribution across relevant subgroups was limited in the COACH study population. Differences in distributions may be more pronounced in another HF population.

The threshold of 70 can be used in designing studies, evaluating interventions, and informing health policy makers.[7]

Conclusion

Scores on the EHFScBs are equally distributed in relevant subgroups of patients with HF included in a trial. For now, it is uncertain whether a clinically relevant change can be defined. A threshold of "70" can be used to discriminate between adequate and inadequate self-care.

Acknowledgments

Author contributions

Disclosure

The authors report no conflicts of interest in this work.

References

1. Jaarsma T, Arestedt KF, Martensson J, Dracup K, Stromberg A. The European Heart Failure Self-care Behaviour scale revised into a nine-item scale (EHFScB-9): a reliable and valid international instrument. *Eur J Heart Fail*. 2009;11(1):99–105.
2. Jaarsma T, Stromberg A, Martensson J, Dracup K. Development and testing of the European Heart Failure Self-Care Behaviour Scale. *Eur J Heart Fail*. 2003;5(3):363–370.
3. Vellone E, Jaarsma T, Stromberg A, et al. The European Heart Failure Self-care Behaviour Scale: new insights into factorial structure, reliability, precision and scoring procedure. *Patient Educ Couns*. 2014; 94(1):97–102.
4. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737–745.
5. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine*. Cambridge: Cambridge University Press; 2011.
6. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol*. 2010;10:22.
7. Thompson LE, Bekelman DB, Allen LA, Peterson PN. Patient-reported outcomes in heart failure: existing measures and future uses. *Curr Heart Fail Rep*. 2015;12(3):236–246.
8. Jaarsma T, Van Der Wal MH, Hogenhuis J, et al. Design and methodology of the COACH study: a multicenter randomised coordinating study evaluating outcomes of advising and counselling in heart failure. *Eur J Heart Fail*. 2004;6(2):227–233.
9. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol*. 2006;59(10):1087–1091.
10. Jaarsma T, van der Wal MH, Lesman-Leegte I, et al. Effect of moderate or intensive disease management program on outcome in patients with heart failure: coordinating study evaluating outcomes of advising and counseling in heart failure (COACH). *Arch Intern Med*. 2008;168(3):316–324.
11. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR; Clinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc*. 2002; 77(4):371–383.
12. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102–109.
13. Brozek JL, Guyatt GH, Schunemann HJ. How a well-grounded minimal important difference can enhance transparency of labelling claims and improve interpretation of a patient reported outcome measure. *Health Qual Life Outcomes*. 2006;4:69.
14. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003; 56(5):395–407.
15. Riegel B, Lee CS, Dickson VV, Carlson B. An update on the self-care of heart failure index. *J Cardiovasc Nurs*. 2009;24(6):485–497.
16. Riegel B, Lee CS, Dickson VV; Medscape. Self care in patients with chronic heart failure. *Nat Rev Cardiol*. 2011;8(11):644–654.
17. Guyatt GH. Measurement of health-related quality of life in heart failure. *J Am Coll Cardiol*. 1993;22(4 suppl A):185A–191A.
18. Taylor AL, Ziesche S, Yancy C, et al. Combination of isosorbide dinitrate and hydralazine in blacks with heart failure. *N Engl J Med*. 2004;351(20):2049–2057.
19. Lee CS, Moser DK, Lennie TA, Tkacs NC, Margulies KB, Riegel B. Biomarkers of myocardial stress and systemic inflammation in patients who engage in heart failure self-care management. *J Cardiovasc Nurs*. 2011;26(4):321–328.
20. Maisel A, Mueller C, Adams K Jr, et al. State of the art: using natriuretic peptide levels in clinical practice. *Eur J Heart Fail*. 2008;10(9): 824–839.
21. Wyrwich KW, Norquist JM, Lenderking WR, Acaster S; Industry Advisory Committee of International Society for Quality of Life Research (ISOQOL). Methods for interpreting change over time in patient-reported outcome measures. *Qual Life Res*. 2013;22(3): 475–483.
22. Riegel B, Dickson VV, Faulkner KM. The situation-specific theory of heart failure self-care: revised and updated. *J Cardiovasc Nurs*. 2016; 31(3):226–235.
23. Bijkerk CJ, Muris JW, Knottnerus JA, Hoes AW, de Wit NJ. Randomized patients in IBS research had different disease characteristics compared to eligible and recruited patients. *J Clin Epidemiol*. 2008; 61(11):1176–1181.
24. Wang TS, Hellkamp AS, Patel CB, Ezekowitz JA, Fonarow GC, Hernandez AF. Representativeness of RELAX-AHF clinical trial population in acute heart failure. *Circ Cardiovasc Qual Outcomes*. 2014; 7(2):259–268.
25. McAlister FA, Stewart S, Ferrua S, McMurray JJ. Multidisciplinary strategies for the management of heart failure patients at high risk for admission: a systematic review of randomized trials. *J Am Coll Cardiol*. 2004;44(4):810–819.
26. Packer M. Proposal for a new clinical end point to evaluate the efficacy of drugs and devices in the treatment of chronic heart failure. *J Card Fail*. 2001;7(2):176–182.
27. Shrive FM, Stuart H, Quan H, Ghali WA. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. *BMC Med Res Methodol*. 2006;6:57.
28. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338:b2393.
29. Evangelista LS, Berg J, Dracup K. Relationship between psychosocial variables and compliance in patients with heart failure. *Heart Lung*. 2001;30(4):294–301.