Original Article

OPEN

Development and Validation of a Short Version of the Kansas City Cardiomyopathy Questionnaire

John A. Spertus, MD, MPH; Philip G. Jones, MS

Background—There is a growing demand to collect patients' experiences of their health status (their symptoms, function, and quality of life) in clinical trials, quality assessment initiatives, and in routine clinical care. In heart failure, the 23-item, disease-specific Kansas City Cardiomyopathy Questionnaire (KCCQ) has been shown to be valid, reliable, sensitive to clinical change, and prognostic of both clinical events and costs. However, its use has been limited, in part, by its length. We sought to develop a shortened version of the instrument that maintains the psychometric properties of the full KCCQ.

Methods and Results—Using data from 3 clinical studies incorporating 4168 patients, we derived and validated a 12-item KCCQ, the KCCQ-12, to capture symptom frequency, physical and social limitations, and quality of life impairment as a result of heart failure, as well as an overall summary score. The KCCQ-12 scores had high correlations with the original scales (>0.93 for all scales in all clinical settings), high test–retest reliability (>0.76 for all domains), high responsiveness (16–31 point improvements after discharge from hospitalization; standardized response mean =0.61–1.12), and comparable prognostic significance and interpretation of clinically important differences as compared with the full KCCQ.

Conclusions—The KCCQ-12 is a shorter version of the original 23-item instrument that should be more feasible to implement while preserving the psychometric properties of the full instrument. (*Circ Cardiovasc Qual Outcomes*. 2015;8:469-476. DOI: 10.1161/CIRCOUTCOMES.115.001958.)

Key Words: heart failure ■ methodology ■ patient-reported outcomes ■ quality ■ quality assessment ■ quality of life

Quantifying patients' perspectives about the degree to which their heart failure (HF) impacts their health status (their symptoms, function, and quality of life) is becoming an increasingly important outcome in clinical trials, quality assessment, and clinical care.[1] Accordingly, over the past 3 decades, several disease-specific patient-reported outcomes (PROs) for HF have been created.[2–6] Transitioning PROs from outcomes in clinical trials, where the studies pay for additional data collection, to routine clinical care requires that the measures be short and feasible to collect, while also retaining the important psychometric properties of validity, reliability, sensitivity to clinical change, prognostic importance, and interpretability.[7,8] The Kansas City Cardiomyopathy Questionnaire (KCCQ), a commonly used instrument for measuring health status in patients with HF, has excellent psychometric properties,[2,9–13] but is currently 23 items long, taking 5 to 8 minutes for patients to complete. Creating shorter and simpler health status measures is a critical step in supporting their use in clinical care[14,15] and can support inexpensive, serial monitoring of patients' HF that might help identify those who warrant additional testing with biomarkers[16,17] or physiological measures[18–20] or treatment intensification.

Given the importance of being able to accurately and objectively assess patients' health status and prognosis using a low-cost, noninvasive strategy, we sought to develop a shorter version of the KCCQ that preserves the psychometric and prognostic properties of the original instrument. Creating a shortened KCCQ meets an important gap in current assessment methods of patients' health status by enabling a more feasible strategy for serially collecting, quantifying, and monitoring the health of HF patients. Accordingly, using KCCQ data from several existing studies of HF patients that have previously demonstrated the psychometric properties of the instrument, we developed and validated a shortened 12-item KCCQ (KCCQ-12). This report describes the development and validation of the KCCQ-12, including its psychometric properties (validity, reproducibility, and responsiveness) and interpretability.

Methods

The Kansas City Cardiomyopathy Questionnaire

The 23-item KCCQ quantifies 7 domains of patients' HF-related health status: Physical Limitation (6 items), Symptom Stability (1 item), Symptom Frequency (4 items), Symptom Burden (3 items),

Self-Efficacy (2 items), Quality of Life (3 items), and Social Limitations (4 items). Item responses are coded sequentially (1, 2, 3, etc.) from worst to best status. Scores are generated for each domain and scaled from 0 to 100, with 0 denoting the worst and 100 the best possible status. In addition, several summary scores are calculated: a Total Symptom score (average of Symptom Frequency and Symptom Burden), a Clinical Summary score (average of Physical Limitation and Total Symptoms), and an Overall Summary score (average of Physical Limitation, Total Symptoms, Quality of Life, and Social Limitation).[2] The KCCQ has been shown to be valid, reproducible, and sensitive to clinical change in patients with systolic dysfunction, HF with preserved ejection fraction, and valvular heart disease.[2,13,21,22] Moreover, patients' KCCQ scores are independently prognostic of survival, HF admissions, and costs.[9–12]

Data Sources

To create a shorter KCCQ, we used data from 2 randomized trials and one observational cohort study of HF patients, all of which have been previously used to demonstrate some of the psychometric properties of the full instrument: (1) the Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study With Tolvaptan (EVEREST) study, a randomized trial of patients hospitalized with HF;[23] (2) the Eplerenone Post-AMI Heart Failure Efficacy and Survival (EPHESUS) study, a randomized trial of patients with HF complicating acute myocardial infarction;[24] and (3) the KCCQ Interpretability Study (KCCQINT), a 14-center North American cohort study of patients presenting at outpatient HF clinics.[13] Descriptions of the studies and KCCQ samples used in the various analyses are provided in Figure. Using these studies, we derived and validated the short KCCQ within 3 distinct clinical settings: (1) stable HF, (2) outpatient HF clinic visits, and (3) acute HF recovery (1 week after hospitalization for decompensated HF). Within each setting, data were split randomly into 2 50% samples, one for derivation (ie, item selection) and one for validation of the final short-version scores. All studies underwent IRB approval before their conduct, and each patient signed informed consent to participate.

Derivation of the Short KCCQ

In deriving the short KCCQ, we first restricted consideration to domains that directly measure patients' current health status: physical limitation, symptom frequency, quality of life, and social limitation. We did not include the single-item symptom change scale because this could not be reduced further, is not incorporated into any of the summary scores, and could always be added to the short version of the KCCQ in clinical settings where a more responsive assessment of recent changes in patients' symptoms was desired. We excluded the self-efficacy scale because it measures a distinctly different concept than the ways in which HF impacts patients' health status. We also excluded the 3 Symptom Burden items, which ask patients how bothersome their symptoms (edema, fatigue, dyspnea) are because they were each highly correlated with their corresponding Symptom Frequency item (r=0.78–0.87) and had less response variability.

For the Symptom Frequency domain, we chose to retain all 4 items to fully represent the spectrum of HF symptoms. For each of the 3 remaining domains, we sought to select items that would maximize comparability between the short and full versions of their domain score. To accomplish this, we examined multiple versions constructed by selecting different possible subsets of items. For the physical limitation domain, which covers low, moderate, and high intensity activities (2 items each), we sought to preserve the range of activities represented by selecting one item from each level of exertional demand, resulting in 8 possible subsets. For the quality of life and social limitation domains, we considered all possible subsets of items (omitting the "intimate relationships" item from the latter domain as a result of high nonresponse rates), yielding 6 and 7 subsets, respectively. For each subset of items, we then calculated a short-version domain score
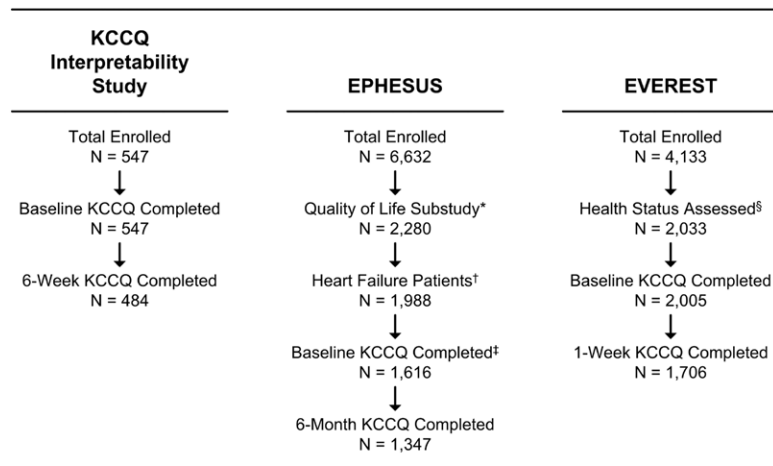
Figure. Data sources and study samples. EPHESUS indicates Eplerenone Post-Acute Myocardial Infarction Heart Failure Efficacy and Survival; EVEREST, Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study With Tolvaptan; and KCCQ, Kansas City Cardiomyopathy Questionnaire.

*Quality of life substudy conducted only in selected countries.
†Patients with diabetes only (no clinical heart failure) excluded.
‡KCCQ assessments completed only on a subset of patients.
§Health status assessments introduced midway through the trial.

following the KCCQ scoring methodology and examined how closely the short score tracked with the corresponding full-version score. We used Lin's concordance correlation coefficient, which measures the agreement between 2 variables; concordance values range from −1 (perfect negative agreement) to 1 (perfect positive agreement), with 0 denoting no agreement.[25] Subsets whose scores had the highest concordance with the full-version score were preferred. Analyses were conducted for each of the 3 clinical settings described above (stable HF, outpatient HF clinic visit, acute HF recovery). Item response variability, nonresponse rates, and clinical judgment were also considered in selecting items.

Once the final set of items was identified, scores for each of the 4 domains were calculated using methodology analogous to that of the full KCCQ, so that scores ranged from 0 to 100 for each domain. In addition, an overall summary KCCQ score was derived as the average of the 4 domain scores, as in the full KCCQ.

Validation

Across the 3 clinical settings, we conducted a series of analyses in the independent, validation samples to evaluate construct validity, predictive validity, reproducibility, and responsiveness and to calculate minimal clinically important differences for the short KCCQ scores. Parallel analyses were conducted for the full KCCQ for comparison. Table 1 summarizes the specific analyses performed and the cohorts used for each analysis.

Construct Validity

To evaluate construct validity, we first compared each short KCCQ domain scores with their respective score from the full KCCQ. Means and standard deviations of scores, as well as mean and standard deviation of differences and concordance coefficients, as described earlier, were calculated and reported. In addition, we calculated mean KCCQ Overall Summary scores by New York Heart Association class I to IV and estimated the association using Kendall's τ-b rank correlation coefficient.

Predictive Validity

Predictive validity refers to the association of scores with subsequent clinical outcomes. For these analyses, the KCCQ scores were used to predict the outcome of 6-month death or cardiovascular hospitalization among stable HF patients (EPHESUS Month 6 data) and among

patients recovering from acute HF hospitalization (EVEREST Week 1 data). Cumulative 6-month incidence was calculated using Kaplan–Meier methods within predefined score categories of 0–<25 (poor), 25–<50 (fair), 50–<75 (good), and 75 to 100 (excellent), as used in previous studies.[9–12] Discrimination was assessed by c-statistics.[26]

Reproducibility

Reproducibility, or test–retest reliability, assesses the change in scores for patients whose clinical status has not changed. We assessed the reproducibility of the short KCCQ by comparing baseline and 6-week change among HF clinic patients (KCCQINT) who had no clinical change based on both patient and physician global health assessments and who had no intervening clinical events.[13] We calculated the mean and standard deviation of 6-week change scores as well as intraclass correlations. The intraclass correlation denotes the proportion of variability in scores because of between-patient (versus within-patient) differences; intraclass correlations >0.4, 0.6, and 0.8 indicate moderate, substantial, and excellent reproducibility.[27]

Responsiveness

Responsiveness is the converse of reproducibility and assesses the sensitivity of the measure to clinical change.[28,29] The responsiveness of the short KCCQ to clinical change was quantified by the change from baseline to 1 week after hospitalization for acute HF (EVEREST). We calculated the mean and standard deviation of change as well as the standardized response mean, defined as the mean change divided by the standard deviation of change. Standardized response means above 0.5 and 0.8 indicate moderate and strong responsiveness, respectively.[27]

Minimal Clinically Important Difference

A critically important issue in the use of health status measures is to define what magnitude of change is clinically important.[30,31] In the KCCQ Interpretability Study, physicians were asked to complete a global assessment of patients' clinical change from baseline, using a 15-point scale ranging from −7 ("a great deal more limited") to +7 ("a great deal less limited") at the follow-up visit.[13,31] In fact, this study was designed explicitly to determine the average change in KCCQ scores across different magnitudes of clinical change, and the clinicians were focused on carefully assessing, blinded to patients' KCCQ scores, the changes in patients' HF status. We used these assessments

Table 1.   Analyses, Settings, and Cohorts Used for Deriving and Validating the KCCQ-12

| Objective	Analyses Performed	Clinical Settings Analyzed		
		Stable HF (EPHESUS Month 6)	HF Clinic Visit (KCCQINT Baseline)	Acute HF Recovery (EVEREST Week 1)
Item selection	Mean±SD; percent missing; concordance with full score	✓	✓	✓
Score validation				
Descriptive statistics	Mean±SD; percent missing	✓	✓	✓
Construct validity	Mean±SD of difference; concordance with full score	✓	✓	✓
	Association with NYHA class (mean±SD; R^2); correlation with 6-minute walk test; correlation with EQ-5D	✓	✓	✓
Predictive validity	6-month death or cardiovascular hospitalization (Kaplan–Meier estimates; c-statistic)	✓		✓
Reproducibility	6-week change in stable patients (mean±SD; intraclass correlation)		✓	
Responsiveness	1-week change after acute HF hospitalization (mean±SD; standardized response mean)			✓
Interpretability	6-week change by physician global assessment categories (mean±SD)		✓	

EPHESUS indicates Eplerenone Post-Acute Myocardial Infarction Heart Failure Efficacy and Survival; EVEREST, Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study With Tolvaptan; HF, heart failure; KCCQ, Kansas City Cardiomyopathy Questionnaire; KCCQINT, KCCQ Interpretability Study; NYHA, New York Heart Association; and SD, standard deviation.

Table 2.   Item Descriptive Statistics: Derivation Samples

	Stable HF (N=673)		HF Clinic Visit (N=273)		Acute HF Recovery (N=853)	
	Mean±SD*	Missing (%)	Mean±SD*	Missing (%)	Mean±SD*	Missing (%)
Physical limitation						
Low intensity						
1a. Dressing yourself	4.6±0.9	4	4.4±0.9	1	4.0±1.1	2
1b. Showering/bathing	4.6±0.9	4	4.4±1.0	0	3.9±1.2	2
Moderate intensity						
1c. Walking 1 block	4.2±1.2	5	3.6±1.4	2	3.4±1.3	4
1d. Yard work/housework	3.8±1.3	8	3.1±1.4	3	2.7±1.3	11
High intensity						
1e. Climbing stairs	3.8±1.3	7	3.0±1.5	3	2.7±1.3	7
1f. Hurrying/jogging	3.0±1.5	15	2.1±1.4	7	1.9±1.1	12
Symptom frequency						
3. Swelling frequency	4.6±1.0	2	4.0±1.4	1	3.6±1.4	1
5. Fatigue frequency*	5.3±1.8	1	4.1±2.0	1	3.8±1.9	1
7. Dyspnea frequency*	5.6±1.8	1	4.4±2.1	1	4.1±2.0	1
9. Dyspnea – sleeping upright	4.6±1.0	1	4.0±1.4	0	4.0±1.4	1
Quality of Life						
12. Enjoyment of life	4.0±1.1	1	3.3±1.3	0	3.1±1.2	1
13. Rest of life as is now	3.5±1.2	1	2.9±1.3	0	2.6±1.2	0
14. Discouraged or down	4.0±1.1	1	3.5±1.1	1	3.4±1.2	1
Social limitation						
15a. Hobbies/recreation	3.8±1.3	9	3.0±1.4	8	2.9±1.3	15
15b. Working/chores	3.9±1.4	25	3.0±1.3	5	2.9±1.3	10
15c. Visiting family/friends	4.2±1.1	7	3.8±1.2	8	3.1±1.4	12
15d. Intimate relationships	3.8±1.3	9	3.0±1.5	22	2.7±1.5	37

HF indicates heart failure; and KCCQ, Kansas City Cardiomyopathy Questionnaire.

*Response range is 1–5 for all items except No. 5 (fatigue frequency) and No. 7 (SOB frequency), which are 1–7. Responses of 6 for physical limitation items (Limited for other reasons or did not do the activity) and social limitation items (Does not apply or did not do for other reasons) are treated as missing, per the KCCQ scoring algorithm.

to determine the minimal clinically important difference for the short KCCQ Summary Score. ROC analyses were conducted predicting any significant improvement (global assessment rating ≥2) and any significant deterioration (≤-2) on the basis of the 6-week change in KCCQ scores. For each end point, the optimal KCCQ cut point was chosen at the point maximizing Youden's Index, which weights sensitivity and specificity equally.[32] Cut point confidence intervals were derived using bootstrap methods. Given prior reports demonstrating that mean group differences in KCCQ scores ≥5 points are clinically important,[13] we also report the sensitivity and specificity of intraindividual changes in patients' scores of these magnitude being clinically important. These analyses were conducted on the entire KCCQ Interpretability Study population to fully use all available data.

Results

Across the 3 studies represented in these analyses, KCCQ data were available on 4168 patients. Descriptions of the studies and the psychometric assessments performed within each study are listed in Table 1.

Derivation

Item selection was conducted within independent derivation samples (50% of the available data, selected randomly) representing each of the 3 clinical settings: stable HF recovering from an acute myocardial infarction (EPHESUS 6-month

assessment, N=673), HF outpatient clinic visit (KCCQ Interpretability Study baseline assessment, N=273), and recovery from acute HF hospitalization (EVEREST Week 1 assessment, N=853). Item response means, standard deviations, and missing rates within each of these settings are outlined in Table 2. In general, stable HF patients had few limitations, minimal symptoms, and good quality of life; in comparison, HF clinic patients had slightly more symptoms and worse quality of life, and patients recovering from HF hospitalization had the worst health status across the 3 domains. Nonresponse rates were low, in general, although higher for physical limitations hurrying or jogging and for all social limitation items (Table 2).

From the 6-item physical limitation domain, concordance with the full score was excellent for all possible 3-item subsets, including one item from each activity level (Appendix I in the Data Supplement). We selected items 1b (limitation with showering or bathing), 1c (limitation with gardening, vacuuming, and carrying groceries), and 1f (limitation hurrying or jogging) as representing a good balance of clinical relevance, item variability, and response rates. From the symptom frequency scale, we retained all 4 items to preserve the varied manifestations of HF symptoms. From the 3-item quality of life scale, concordance was excellent for all 2-item subsets and superior to single-item

subsets. We selected items 12 (limitation of enjoyment of life) and 13 (feelings about spending the rest of your life with symptoms as they are now) for clinical relevance and item variability. Finally, from the social limitation scale, we omitted item 15d (limitation in intimate relationships) because of high nonresponse rates. We retained all 3 remaining items because that yielded the highest concordance with the full social limitation score among all possible subsets. In summary, we identified 12 items (3 physical limitation, 4 symptom frequency, and 2 quality of life and 3 social limitation items) to retain in the final short version of the KCCQ (Appendix II in the Data Supplement).

Validation

Construct Validity
Agreement between the KCCQ-12 and full KCCQ scores was excellent in all clinical settings, with concordances of 0.97 for physical limitation scores, 0.93 to 0.96 for quality of life scores, 0.98 for social limitation scores, and 0.98 to 0.99 for summary scores (Appendix III in the Data Supplement). Concordance was perfect for the symptom frequency domain because the same items are used in both instruments to generate that domain score. As with the full KCCQ, missing data occurred primarily for physical limitation and social limitation scores and was slightly more frequent in the KCCQ-12, but ≤10% in all clinical settings. The KCCQ-12 summary score also demonstrated a strong association with New York Heart Association class, comparable to that of the full KCCQ, with mean±SD summary scores ranging from 29±29 for Class IV to 86±15 for Class I patients with stable CAD ($r=-0.43$) and from 26±17 for Class IV to 70±22 for Class I patients after hospitalization for HF ($r=-0.35$; Table 3).

Predictive Validity
All KCCQ-12 scores demonstrated a graded inverse relationship with the outcome of 6-month death or cardiovascular (CV) hospitalization, comparable to those of the full KCCQ (Table 4). Predictive power was consistently strongest for the overall summary score, with cumulative incidence among stable HF patients ranging from 7% for those with scores 75 to 100 to 33% for those with scores <25 (c-index=0.64) and among patients recovering from HF hospitalization ranging from 24% for those with scores 75 to 100 to 62% for those with scores <25 (c-index=0.63).

Reproducibility
Among 79 clinically stable patients assessed 6 weeks after a HF clinic visit, the KCCQ-12 revealed minimal changes in

scores between assessments and showed high intraclass correlations of ≥0.76 for all domains (Table 5). The overall summary score had the highest reproducibility with an intraclass correlation of 0.92.

Responsiveness
The KCCQ-12 showed substantial responsiveness to clinical change. One week after hospitalization for acute HF, mean KCCQ-12 scores increased by >16 points for all domains, with the greatest increase in the symptom frequency score (31.0 points; Table 6). Standardized response means were good to excellent, ranging from 0.65 for physical limitation to 1.12 for the summary score.

Minimal Clinically Important Difference
Among patients at 6 weeks after a HF clinic visit, change in the KCCQ-12 summary score was strongly associated with physicians' assessment of clinically significant improvement (c-index =0.67) and deterioration (c-index =0.77), mirroring that of the full KCCQ (c-indices of 0.68 and 0.77, respectively). The optimal cut point based on Youden's index was 4.7 points for predicting improvement and −3.1 points for predicting deterioration, both agreeing favorably with analogous cut points for the full KCCQ (Table 7). A 5-point improvement or deterioration in scores for individual patients was associated with a 58% and 67% sensitivity and a 70% and 75% specificity, respectively, that a physician would have assessed the patient to have a clinically important improvement or deterioration in their HF status. These results were comparable to the full KCCQ and support a minimal clinically important difference of ≈3 to 5 points for the KCCQ-12 summary score.

Discussion
As the US healthcare system struggles to provide more cost-effective, patient-centered care, there is a growing need to efficiently monitor the health status of its patients.[33] For example, the Centers for Medicare and Medicaid Services has begun soliciting proposals for PROs measures to become part of its quality assessment programs (http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/CallForMeasures.html). A critical challenge in using PROs for quality assessment is that they must be feasible to collect and incorporate into routine clinical care so that there is minimal burden on practitioners for reporting patients' health status to payers. Moreover, creating a shortened version of the KCCQ provides a more feasible means

Table 3.   Association Between KCCQ Summary Scores and NYHA Class

	Stable HF (N=674)			HF Clinic Visit (N=274)			Acute HF Recovery (N=853)		
	N	KCCQ-12	Full KCCQ	N	KCCQ-12	Full KCCQ	N	KCCQ-12	Full KCCQ
NYHA class									
I	319	86±15	86±15	29	80±15	81±15	40	70±22	71±22
II	267	73±19	73±19	115	69±20	70±20	366	58±20	60±20
III	80	50±21	51±21	115	51±22	52±21	387	45±19	47±18
IV	7	29±29	31±27	14	28±25	30±24	55	26±17	28±16
Correlation		−0.43	−0.44		−0.41	−0.41		−0.35	−0.34

HF indicates heart failure; KCCQ, Kansas City Cardiomyopathy Questionnaire; and NYHA, New York Heart Association.

Table 4.   6-Month Death or Cardiovascular Hospitalization by KCCQ Score

	Stable HF (N=674)		Acute HF Recovery (N=853)	
	KCCQ-12	Full KCCQ	KCCQ-12	Full KCCQ
Physical limitation				
Poor (0–<25)	18%±7%	15%±6%	51%±5%	48%±5%
Fair (25–<50)	12%±4%	10%±4%	43%±3%	43%±3%
Good (50–<75)	12%±3%	12%±3%	31%±3%	30%±3%
Excellent (75–100)	6%±1%	6%±1%	25%±3%	24%±3%
C-statistic	0.61	0.61	0.59	0.59
Symptom frequency				
Poor (0–<25)	10%±9%	10%±9%	57%±5%	57%±5%
Fair (25–<50)	13%±5%	13%±5%	42%±4%	42%±4%
Good (50–<75)	13%±3%	13%±3%	34%±3%	34%±3%
Excellent (75–100)	7%±1%	7%±1%	27%±3%	27%±3%
C-statistic	0.61	0.61	0.59	0.59
Quality of life				
Poor (0–<25)	23%±7%	25%±8%	58%±5%	53%±5%
Fair (25–<50)	14%±4%	13%±4%	33%±3%	42%±3%
Good (50–<75)	8%±2%	9%±2%	33%±3%	30%±3%
Excellent (75–100)	6%±1%	6%±1%	28%±4%	27%±4%
C-statistic	0.63	0.62	0.58	0.58
Social limitation				
Poor (0–<25)	14%±6%	13%±5%	56%±4%	56%±4%
Fair (25–<50)	16%±5%	19%±5%	38%±3%	34%±3%
Good (50–<75)	9%±3%	7%±2%	30%±3%	33%±4%
Excellent (75–100)	6%±1%	6%±1%	22%±3%	20%±3%
C-statistic	0.60	0.62	0.60	0.60
Summary score				
Poor (0–<25)	33%±11%	31%±12%	62%±5%	62%±5%
Fair (25–<50)	14%±4%	11%±4%	38%±3%	41%±3%
Good (50–<75)	7%±2%	11%±2%	30%±3%	29%±3%
Excellent (75–100)	7%±1%	6%±1%	24%±4%	25%±4%
C-statistic	0.63	0.63	0.61	0.61

HF indicates heart failure; and KCCQ, Kansas City Cardiomyopathy Questionnaire.

to routinely assess patients' health status to improve clinical care and population management. In this study, we have been able to substantially reduce the KCCQ from 23 to 12 items, while retaining the validity, reproducibility, responsiveness, predictive validity, and interpretability of the original instrument. Moreover, the high concordance (0.93–1.0) between the domain scores of the 2 versions—established in several distinct clinical circumstances—suggests that the scores can be used interchangeably.

The availability of a shortened disease-specific health status measure for patients with HF has the potential to improve care. First, quantifying patients' health status at each outpatient visit has been endorsed as a performance measure of healthcare quality.[34] Although this measure can be met with either the New York Heart Association classification or 1 of 3 PROs, including the KCCQ, the inter-rater reliability of the New York Heart Association is poor,[35–37] with a concordance of 54% as compared with the intraclass correlation coefficient of 92% for the KCCQ-12 overall summary score. Having

more reliable estimates of patients' health status through self-reported PROs becomes, in essence, a standardized history that can provide a better estimate from which to assess whether patients' conditions have changed over time. In addition, incorporating a short PRO into the clinical examination can theoretically improve the efficiency of a provider's visit

Table 5.   6-Week Change in Stable Patients (N=79)

	KCCQ-12		Full KCCQ	
	Mean±SD	Intra-Class Correlation	Mean±SD	Intra-Class Correlation
Physical limitation	2.8±14.4	0.85	2.4±13.5	0.86
Symptom frequency	1.0±14.5	0.83	1.0±14.5	0.83
Quality of life	3.4±19.8	0.76	2.4±16.2	0.82
Social limitation	3.8±16.0	0.86	4.0±16.4	0.85
Summary score	2.6±10.5	0.91	2.3±9.7	0.92

KCCQ indicates Kansas City Cardiomyopathy Questionnaire.

Table 6. 1-Week Change After Acute Heart Failure Hospitalization (N=853)

	KCCQ-12		Full KCCQ	
	Mean±SD	Standardized Response Mean	Mean±SD	Standardized Response Mean
Physical limitation	16.5±25.2	0.65	17.2±24.2	0.71
Symptom frequency	31.0±27.5	1.12	31.0±27.5	1.12
Quality of life	21.6±26.1	0.83	20.2±24.1	0.84
Social limitation	17.6±28.7	0.61	17.2±27.8	0.62
Summary score	21.8±21.3	1.02	21.1±20.8	1.02

KCCQ indicates Kansas City Cardiomyopathy Questionnaire.

Table 7. Minimal Clinically Important Differences in KCCQ Overall Summary Score by 6-Week Physician Global Assessment (N=547)

	KCCQ-12	Full KCCQ
Any clinically significant improvement (n=103)		
Youden point (95% CI)	4.7 (−6.3, 7.8)	3.6 (2.1, 6.1)
Sensitivity	60%	67%
Specificity	69%	67%
Overall summary score ≥5 points		
Sensitivity	58%	58%
Specificity	70%	72%
Any clinically significant deterioration (n=52)		
Youden point (95% CI)	−3.1 (−6.9, −1.0)	−3.6 (−10.2, −1.8)
Sensitivity	75%	69%
Specificity	69%	74%
Overall summary score ≤−5 points		
Sensitivity	67%	58%
Specificity	75%	78%

CI indicates confidence interval; and KCCQ, Kansas City Cardiomyopathy Questionnaire.

by enabling providers to know, at a glance, whether the patient is doing better or worse as compared with the prior visit. We recently cited this as a potential benefit of the routine use of PROs and highlighted the need for a feasible version of PROs to support their adoption in clinical care.[15] In addition, the foundation of population health and disease management programs is to assess the health status of many patients and to direct resources to those with the greatest potential to benefit. Were health systems to assess the health status of their HF patients with the KCCQ-12, they could readily identify those with the worst symptoms and highest risk for mortality and hospitalization[10–12,22] and those expected to cost the most,[9] so that they could preferentially direct more care to these patients to improve their outcomes. Testing these applications of PROs in clinical practice is a high research priority as healthcare moves from a volume-based to a value-based reimbursement model.

Although there are several disease-specific PROs available for patients with HF, the original KCCQ had the greatest amount of psychometric data supporting its use, including standards for defining clinically important differences in scores,[13] and separately quantified a broader range of clinically important domains, including symptoms, physical limitations, social limitations, and quality of life, than other measures. Accordingly, several recent reviews of PROs in HF have recommended the KCCQ or Minnesota Living with Heart Failure questionnaires as the preferred instruments.[5,6] Creating a shorter version of the KCCQ that maintains all of the psychometric properties of the original instrument should further enhance its value and encourage other instrument developers to explore reducing the length of their measures.

The development of the KCCQ-12, a shortened version of the original 23-item KCCQ, should be interpreted in the context of the following potential limitations. First, although we demonstrated excellent psychometric performance of the KCCQ-12, all of the limitations of the original instrument likely apply to the shortened version. For example, patients who do not perform activities because of conditions other than their HF will have missing scores for the physical limitation domain. Moreover, although the shortened version should minimize missing data, any informatively missing health status assessments need to be handled carefully, through appropriate statistical methods, to avoid introducing bias. In addition, despite data showing comparable performance of the

KCCQ in patients with valvular heart disease or HF because of preserved ejection fraction,[21,22] all of these analyses were performed in patients with reduced ejection fraction, and replication of these analyses in patients with other etiologies of HF may be important.

In conclusion, we have validated a 12-item version of the KCCQ and shown it to have comparable psychometric properties as compared with the original, 23-item version. It preserves the high test–retest reliability, responsiveness, prognostic ability, and interpretability thresholds of the original KCCQ, while reducing the response burden by half. It thus has the potential to be more implementable in routine clinical care and can open up important opportunities to improve the quality of HF care through quality assessment efforts, tailored population health interventions, more practical clinical trials, and better doctor–patient communication. Testing these applications are important opportunities for future research and, if successful, can support more patient-centered care.[33]

Disclosures

References

1. Spertus JA. Evolving applications for patient-centered health status measures. *Circulation*. 2008;118:2103–2110. doi: 10.1161/CIRCULATIONAHA.107.747568.

2. 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.

3. Guyatt GH, Nogradi S, Halcrow S, Singer J, Sullivan MJ, Fallen EL. Development and testing of a new measure of health status for clinical trials in heart failure. *J Gen Intern Med*. 1989;4:101–107.

4. Rector TS, Kubo SH, Cohn JN. Validity of the Minnesota Living with Heart Failure questionnaire as a measure of therapeutic response to enalapril or placebo. *Am J Cardiol*. 1993;71:1106–1107.

5. Garin O, Herdman M, Vilagut G, Ferrer M, Ribera A, Rajmil L, Valderas JM, Guillemin F, Revicki D, Alonso J. Assessing health-related quality of life in patients with heart failure: a systematic, standardized comparison of available measures. *Heart Fail Rev*. 2014;19:359–367. doi: 10.1007/s10741-013-9394-7.

6. Mackintosh A, Gibbons E, Fitzpatrick R. A structured review of patient-reported outcome measures for people with heart failure: an update 2009. http://phi.uhce.ox.ac.uk/pdf/PROMs_Oxford_HeartFailure_17092010.pdf. Accessed July 13, 2015.

7. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med*. 1993;118:622–629.

8. US Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research. Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2006:1–32.

9. Chan PS, Soto G, Jones PG, Nallamothu BK, Zhang Z, Weintraub WS, Spertus JA. Patient health status and costs in heart failure: insights from the eplerenone post-acute myocardial infarction heart failure efficacy and survival study (EPHESUS). *Circulation*. 2009;119:398–407. doi: 10.1161/CIRCULATIONAHA.108.820472.

10. Heidenreich PA, Spertus JA, Jones PG, Weintraub WS, Rumsfeld JS, Rathore SS, Peterson ED, Masoudi FA, Krumholz HM, Havranek EP, Conard MW, Williams RE; Cardiovascular Outcomes Research Consortium. Health status identifies heart failure outpatients at risk for hospitalization or death. *J Am Coll Cardiol*. 2006;47:752–756. doi: 10.1016/j.jacc.2005.11.021.

11. 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.

12. 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.

13. Spertus J, Peterson E, Conard MW, Heidenreich PA, Krumholz HM, Jones P, McCullough PA, Pina I, Tooley J, Weintraub WS, Rumsfeld JS; Cardiovascular Outcomes Research Consortium. Monitoring clinical changes in patients with heart failure: a comparison of methods. *Am Heart J*. 2005;150:707–715. doi: 10.1016/j.ahj.2004.12.010.

14. Rumsfeld JS. Health status and clinical practice: when will they meet? *Circulation*. 2002;106:5–7.

15. Spertus J. Barriers to the use of patient-reported outcomes in clinical care. *Circ Cardiovasc Qual Outcomes*. 2014;7:2–4. doi: 10.1161/CIRCOUTCOMES.113.000829.

16. Schou M, Gustafsson F, Videbaek L, Andersen H, Toft J, Nyvad O, Ryde H, Fog L, Jensen JC, Nielsen OW, Lind-Rasmussen S, Abdulla J, Hildebrandt PR; NorthStar Investigators. Adding serial N-terminal pro brain natriuretic peptide measurements to optimal clinical management in outpatients with systolic heart failure: a multicentre randomized clinical trial (NorthStar monitoring study). *Eur J Heart Fail*. 2013;15:818–827. doi: 10.1093/eurjhf/hft037.

17. Januzzi JL Jr, Rehman SU, Mohammed AA, Bhardwaj A, Barajas L, Barajas J, Kim HN, Baggish AL, Weiner RB, Chen-Tournoux A, Marshall JE, Moore SA, Carlson WD, Lewis GD, Shin J, Sullivan D, Parks K, Wang TJ, Gregory SA, Uthamalingam S, Semigran MJ. Use of amino-terminal pro-B-type natriuretic peptide to guide outpatient therapy of patients with chronic left ventricular systolic dysfunction. *J Am Coll Cardiol*. 2011;58:1881–1889. doi: 10.1016/j.jacc.2011.03.072.

18. Abraham WT. Disease management: remote monitoring in heart failure patients with implantable defibrillators, resynchronization devices, and haemodynamic monitors. *Europace*. 2013;15 Suppl 1:i40–i46. doi: 10.1093/europace/eut105.

19. Loh JP, Barbash IM, Waksman R. Overview of the 2011 Food and Drug Administration Circulatory System Devices Panel of the Medical Devices Advisory Committee Meeting on the CardioMEMS Champion Heart Failure Monitoring System. *J Am Coll Cardiol*. 2013;61:1571–1576. doi: 10.1016/j.jacc.2012.08.035.

20. Abraham WT, Adamson PB, Bourge RC, Aaron MF, Costanzo MR, Stevenson LW, Strickland W, Neelagaru S, Raval N, Krueger S, Weiner S, Shavelle D, Jeffries B, Yadav JS; CHAMPION Trial Study Group. Wireless pulmonary artery haemodynamic monitoring in chronic heart failure: a randomised controlled trial. *Lancet*. 2011;377:658–666. doi: 10.1016/S0140-6736(11)60101-3.

21. Arnold SV, Spertus JA, Lei Y, Allen KB, Chhatriwalla AK, Leon MB, Smith CR, Reynolds MR, Webb JG, Svensson LG, Cohen DJ. Use of the Kansas City Cardiomyopathy Questionnaire for monitoring health status in patients with aortic stenosis. *Circ Heart Fail*. 2013;6:61–67. doi: 10.1161/CIRCHEARTFAILURE.112.970053.

22. Joseph SM, Novak E, Arnold SV, Jones PG, Khattak H, Platts AE, Dávila-Román VG, Mann DL, Spertus JA. Comparable performance of the Kansas City Cardiomyopathy Questionnaire in patients with heart failure with preserved and reduced ejection fraction. *Circ Heart Fail*. 2013;6:1139–1146. doi: 10.1161/CIRCHEARTFAILURE.113.000359.

23. Konstam MA, Gheorghiade M, Burnett JC Jr, Grinfeld L, Maggioni AP, Swedberg K, Udelson JE, Zannad F, Cook T, Ouyang J, Zimmer C, Orlandi C; Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study With Tolvaptan (EVEREST) Investigators. Effects of oral tolvaptan in patients hospitalized for worsening heart failure: the EVEREST Outcome Trial. *JAMA*. 2007;297:1319–1331. doi: 10.1001/jama.297.12.1319.

24. Spertus JA, Tooley J, Jones P, Poston C, Mahoney E, Deedwania P, Hurley S, Pitt B, Weintraub WS. Expanding the outcomes in clinical trials of heart failure: the quality of life and economic components of EPHESUS (EPlerenone's neuroHormonal Efficacy and SUrvival Study). *Am Heart J*. 2002;143:636–642.

25. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45:255–268.

26. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837–845.

27. McDowell I. *Measuring Health: A Guide to Rating Scales and Questionnaires*. New York: Oxford University Press; 2006.

28. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis*. 1987;40:171–178.

29. Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health status measurement: a clarification. *J Clin Epidemiol*. 1989;42:403–408.

30. Guyatt G, Schunemann H. How can quality of life researchers make their work more useful to health workers and their patients? *Qual Life Res*. 2007;16:1097–1105. doi: 10.1007/s11136-007-9223-3.

31. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10:407–415.

32. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3:32–35.

33. Institute of Medicine. *Crossing the Quality Chasm: A New Health System for the Twenty-first Century*. Washington: National Academy Press; 2001.

34. Physician Consortium for Performance Improvement. Clinical Performance Measures in Heart Failure. A Consensus Document from the ACC, AHA and the Consortium. 2006.

35. Bennett JA, Riegel B, Bittner V, Nichols J. Validity and reliability of the NYHA classes for measuring research outcomes in patients with cardiac disease. *Heart Lung*. 2002;31:262–270.

36. Goldman L, Cook EF, Mitchell N, Flatley M, Sherman H, Cohn PF. Pitfalls in the serial assessment of cardiac functional status. How a reduction in "ordinary" activity may reduce the apparent degree of cardiac compromise and give a misleading impression of improvement. *J Chronic Dis*. 1982;35:763–771.

37. Raphael C, Briscoe C, Davies J, Ian Whinnett Z, Manisty C, Sutton R, Mayet J, Francis DP. Limitations of the New York Heart Association functional classification system and self-reported walking distances in chronic heart failure. *Heart*. 2007;93:476–482. doi: 10.1136/hrt.2006.089656.