JACC: HEART FAILURE VOL. 6, NO. 7, 2018

PUBLISHED BY ELSEVIER

CLINICAL RESEARCH

Comparing Measures to Assess Health-Related Quality of Life in Heart Failure With Preserved Ejection Fraction

Rebecca Napier, MD,[a] Steven E. McNulty, MS,[b] David T. Eton, PhD,[c] Margaret M. Redfield, MD,[a] Omar AbouEzzeddine, MD, CM, MS,[a] Shannon M. Dunlay, MD, MS[a,c]

ABSTRACT

OBJECTIVES This study sought to compare the performance of 2 health-related quality of life (HRQOL) questionnaires in patients with heart failure with preserved ejection fraction (HFpEF).

BACKGROUND The ability to accurately assess HRQOL over time is important in the care of patients with heart failure. The validity and reliability of HRQOL tools including the Minnesota Living with Heart Failure Questionnaire (MLHFQ) and the Kansas City Cardiomyopathy Questionnaire (KCCQ) has not been fully determined or compared in patients with HFpEF.

METHODS Among patients with stable chronic HFpEF enrolled in the NEAT (Nitrate Effect on Activity Tolerance in Heart Failure) trial (n = 110), the study evaluated and compared reliability, validity, and responsiveness to change of the MLHFQ and KCCQ at baseline, 6 weeks, and 12 weeks.

RESULTS Internal consistency was good and comparable for MLHFQ and KCCQ domains measuring similar aspects of HRQOL at baseline including the MLHFQ physical (Cronbach's $\alpha = 0.93$) compared with the KCCQ clinical summary ($\alpha = 0.91$), and the MLHFQ emotional ($\alpha = 0.92$) compared with the KCCQ quality of life ($\alpha = 0.87$). Correlations with New York Heart Association functional class (Spearman rho; $r_s = -0.37$ vs. 0.30) and 6-min walk test (6MWT) ($r_s = 0.38$ vs. -0.23) at baseline were slightly stronger for the KCCQ overall summary score than for the MLHFQ total score. The MLHFQ was more responsive to change in 6MWT based on responsiveness statistics.

CONCLUSIONS These data suggest that both the MLHFQ and KCCQ are reliable and valid tools to assess HRQOL in HFpEF. The KCCQ was more strongly correlated with baseline functional status parameters, while the MLHFQ was more responsive to improvement in 6MWT. (J Am Coll Cardiol HF 2018;6:552-60) © 2018 by the American College of Cardiology Foundation.

Heart failure (HF) continues to have an immense impact, not only on tangible outcomes such as mortality (1), but also on patients' perceived health-related quality of life (HRQOL) (2-4). Patients with HF with preserved ejection fraction (HFpEF), who now comprise one-half of all patients with HF (5), struggle with poor HRQOL, similar to patients with HF with reduced ejection

From the [a]Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota; [b]Duke Clinical Research Institute, Durham, North Carolina; and the [c]Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota. This research was supported by grant numbers U10 HL084904 (awarded to the coordinating center) and U01 HL084861, U10HL110312, U10 HL110337, U10 HL110342, U10 HL110262, U10 HL110297, U10 HL110302, U10 HL110309, U10 HL110336, and U10 HL110338 (awarded to the regional clinical centers) from the National Heart, Lung, and Blood Institute. Dr. Dunlay's time is funded by the National Institutes of Health (K23 HL 116643). The authors have reported that they have no relationships relevant to the contents of this paper to disclose.

Manuscript received August 1, 2017; revised manuscript received February 15, 2018, accepted February 18, 2018.

ISSN 2213-1779/$36.00 https://doi.org/10.1016/j.jchf.2018.02.006

fraction (HFrEF) (6). Improving HRQOL remains a primary goal in the management of patients with HFpEF (7), and is reliant on the ability to accurately measure HRQOL when assessing the effectiveness of novel HFpEF therapies.

The Minnesota Living with Heart Failure Questionnaire (MLHFQ) and the Kansas City Cardiomyopathy Questionnaire (KCCQ) are the 2 most widely used instruments to assess disease-specific HRQOL in patients with HF. However, although both instruments have been demonstrated to be reliable and valid in patients with HFrEF (8–10), their performance has not been compared in patients with HFpEF. As HFpEF and HFrEF have distinct pathophysiology, respond differently to therapies, and tend to arise in different patient populations, there may be important differences in HRQOL and its measurement that are important to examine (11).

To provide guidance to clinicians and researchers working with patients with HFpEF, the goal of this study was to evaluate and compare the psychometric performance, including the reliability, validity, and responsiveness to change in clinical status, of the MLHFQ and KCCQ among patients enrolled in the NEAT (Nitrate Effect on Activity Tolerance in Heart Failure) trial (12).

SEE PAGE 561

METHODS

STUDY POPULATION. We included participants enrolled in the NEAT trial. The NEAT trial was a multicenter, double-blind, crossover study conducted over 12 weeks that assessed the effect of isosorbide mononitrate compared with placebo on level of activity in patients with HFpEF (12,13). In symptomatic HFpEF patients, the addition of a long-acting nitrate did not improve activity level, and there was no impact on 6-min walk test (6MWT) distance, HRQOL, or clinical biomarkers. Patients eligible for study enrollment included those 50 years of age and older on stable medical therapy for the treatment of HF with a documented ejection fraction of 50% or greater. After 1:1 randomization to isosorbide mononitrate or placebo (n = 110), data were collected at baseline, 6 weeks, and 12 weeks. All participants underwent crossover at week 6 with each patient receiving both treatments in random order. Treatment was administered for 4 weeks after a 2-week no-drug phase. HRQOL, as assessed by the MLHFQ and the KCCQ, was a secondary endpoint in the trial. As such, the MLHFQ and KCCQ were collected at baseline (visit 1) and after taking graduated doses of isosorbide mononitrate or placebo for 4 weeks after a 2-week no-drug phase (visit 2 after week 6 and visit 3 after week 12). No impact of treatment on HRQOL was observed in the trial. We present the comparison of HRQOL assessments from the baseline and 6-week visits in each patient in the main paper, and have included the 12-week visit data in the Online Appendix.

HRQOL ASSESSMENT TOOLS. The MLHFQ is a 21-item questionnaire with a 6-point response scale ranging from 0 to 5. The questionnaire asks respondents to reflect upon the way they have felt over the last 4 weeks when selecting responses. A total score is obtained by summing the scores from individual questions (range 0 to 105) with higher scores indicating poorer HRQOL (9,10). In addition to the total score, selected items also provide scores specific to physical and emotional HRQOL parameters. A change of ≥5 points is considered clinically meaningful (14).

The KCCQ is a 23-item questionnaire that quantifies multiple HRQOL domains that can be affected in patients with HF, including physical and social limitations, symptom frequency and severity, overall QOL, recent changes in symptom status, and self-efficacy (8). It asks respondents to reflect on the way they have felt over the last 2 weeks when selecting responses. In addition to its individual subscales, there are several combined subscales. The total symptom subscale is the average of symptom frequency and symptom burden. The clinical summary scale is the average of physical limitation and total symptoms. The overall summary score is the average of physical limitations, total symptoms, overall QOL, and social limitations. KCCQ scores range from 1 to 100, with higher scores indicative of better HRQOL. A change of ≥5 points is considered clinically significant (3,15).

STATISTICAL ANALYSIS. Baseline characteristics of the study population were summarized using number (percentage), mean ± SD, and median (interquartile range) where appropriate. Ceiling and floor effects, defined as the top and bottom 5% of possible scores on each instrument, were assessed for the MLHFQ and KCCQ at baseline (visit 1) and after week 6 (visit 2) to determine the extent at which extreme scores may limit the ability of each questionnaire to detect change in HRQOL over time.

Internal consistency and reliability. The internal consistency reliability of the KCCQ and MLHF overall scores and subscales at baseline, 6 weeks, and 12 weeks were assessed using Cronbach's α coefficients. Cronbach's α evaluates the internal consistency, or average intercorrelation, of items within a domain.

ABBREVIATIONS AND ACRONYMS

6MWT = 6-min walk test

HF = heart failure

HFpEF = heart failure with preserved ejection fraction

HFrEF = heart failure with reduced ejection fraction

HRQOL = health-related quality of life

KCCQ = Kansas City Cardiomyopathy Questionnaire

MLHFQ = Minnesota Living with Heart Failure Questionnaire

NT-proBNP = N-terminal pro-B-type natriuretic peptide

NYHA = New York Heart Association

Values range from 0 to 1 with higher values reflecting greater consistency of the items within the scale, with αs ≥0.70 indicating adequate reliability (16). For the MLHFQ, consistencies of items in the physical and emotional subscales were assessed. The KCCQ subscales assessed included the overall summary score, clinical summary score, physical limitation subscale, total symptom subscale, self-efficacy subscale, and overall QOL subscale.

Validity. Criterion validity was assessed by comparing baseline New York Heart Association (NYHA) functional class, 6MWT distance, and N-terminal pro-B-type natriuretic peptide (NT-proBNP) levels, 3 commonly used criterion standards in HF, with MLHFQ and KCCQ scores using Spearman rank-order correlation coefficients (Spearman rho [r_s]). A correlation magnitude of at least moderate level ($r_s \geq 0.30$) and in the hypothesized direction (i.e., better functional class and longer walk distance) would indicate evidence of criterion validity (17). As accelerometry was the primary outcome of the NEAT trial, we also compared the baseline daily average accelerometry units in the 2 weeks before initiation of study drug with MLHFQ and KCCQ scores in an exploratory analysis.

Responsiveness to clinical change. To evaluate the responsiveness to clinical change over time, absolute changes in KCCQ and MLHFQ from baseline to 6 and 12 weeks were correlated with changes in other measures of clinical status including 6MWT and NT-proBNP levels. For the 6MWT, 3 mutually exclusive categories of clinical change were defined based on prior work (2,3): improved (≥+25 m), no change (+24 to −24 m), or deteriorated (≤−25 m). Similarly, NT-proBNP results were categorized as increased (≥20% increase), decreased (≥20% decrease), or unchanged. We chose not to evaluate changes in KCCQ and MLHFQ according to changes in accelerometry over time, as no established cutpoints of clinical change in accelerometry exist. In addition to assessing the responsiveness to change of the MLHFQ total score and KCCQ overall summary score, we selected domains that assessed similar aspects of HRQOL from the KCCQ to compare to the MLHFQ physical (KCCQ clinical summary) and emotional (KCCQ QOL) domains.

Responsiveness parameters evaluated include the *t* statistic (mean change divided by SD for total group), effect size (mean change divided by the SD of the baseline score), Guyatt's responsiveness statistic (mean change divided by the SD of change in subjects who remained unchanged), and the standardized response mean (mean change divided by the SD of the change score). These parameter values were then summarized to determine the overall median rank for each HRQOL score, reflecting the relative responsiveness of each metric to changes in clinical status. All analyses were performed in SAS version 9.4 (SAS Institute, Cary, North Carolina).

RESULTS

Baseline characteristics of the 110 participants with HFpEF in the NEAT trial are shown in **Table 1**. Just over one-half of participants were women (57%), and all but 2 had NYHA functional class II or III symptoms at baseline. The mean MLHFQ score at baseline was 44.6 ± 23.7, whereas the mean KCCQ overall summary score was 56.0 ± 24.2. At baseline, there were 3 (3.2%) patients with MLHFQ total scores of 0 to 5.25 (ceiling), and 2 (2.1%) with scores of 99.75 to 105.00 (floor). For the KCCQ overall summary score, there was 1 (0.9%) patient with a score of 0 to 5 (floor) and 5 (4.6%) patients with scores of 95 to 100 (ceiling). Thus, the ability to detect change over time in HRQOL was not limited by extremes in scoring for either instrument.

In total, 78 patients completed the MLHFQ at both baseline and 6 weeks. By comparison, 108 patients completed the KCCQ at baseline and 6 weeks. A total of 78 patients had complete MLHFQ and KCCQ data at both baseline and 6 weeks. The proportions of participants with improvement (≥5-point improvement), no change, and decline (≥5-point worsening) in HRQOL scores from baseline to week 6 were 55%, 28%, and 17% for the MLHFQ and 44%, 40%, and 16% for the KCCQ. The level of agreement of change in MLHFQ and KCCQ was fair (Cohen's kappa statistic: 0.36; 95% confidence interval: 0.20 to 0.52) (Online Table 1). The MLHFQ (n = 78) and KCCQ (n = 107) were reassessed at 12 weeks. A total of 77 patients completed both the MLHFQ and KCCQ at baseline and 12 weeks. The proportions of patients with improvement, no change, and decline in HRQOL scores from baseline to 12 weeks were similar to 6 weeks (Online Table 2).

INTERNAL CONSISTENCY AND RELIABILITY. Internal consistency reliability of most of the MLHFQ and KCCQ summary and domain scales was adequate. Only the 6-week KCCQ self-efficacy scale had a Cronbach's α <0.70 (**Table 2**).

VALIDITY. The correlations of MLHFQ and KCCQ scores with NYHA functional class, 6MWT distance, NT-proBNP, and average accelerometry units at baseline are shown in **Table 3 and** Online Table 3. The MLHFQ total score, in which higher scores indicate worse HRQOL, was moderately correlated with NYHA functional class ($r_s = 0.30$; $p = 0.003$) and 6MWT ($r_s = -0.23$; $p = 0.024$). Moderate correlations with NYHA functional class ($r_s = 0.35$; $p < 0.001$) and

6MWT distance ($r_s = -0.39$; $p < 0.001$) were also observed for the MLHFQ physical score, whereas the correlations with MLHFQ emotional score were weak and not statistically significant. There was a modest negative correlation between MLHFQ emotional score and NT-proBNP levels ($r_s = -0.21$; $p = 0.027$). This is opposite of the hypothesized direction, and may be spurious. No other significant correlations of HRQOL scores and NT-proBNP were observed. Correlations of the MLHFQ total, physical, and emotional scores with baseline average accelerometry units were in the hypothesized direction, but were weak and not statistically significant.

For the KCCQ (higher scores reflect better HRQOL) there were moderate correlations of the overall summary score ($r_s = -0.37$, $r_s = 0.38$), clinical summary score ($r_s = -0.37$, $r_s = 0.44$), physical limitation score ($r_s = -0.38$, $r_s = 0.47$), and total symptoms score ($r_s = -0.30$, $r_s = 0.31$), with NYHA functional class and 6MWT, respectively ($p < 0.001$ for each). The KCCQ QOL score was weakly correlated with NYHA functional class ($r_s = -0.28$; $p = 0.003$) and 6MWT ($r_s = 0.19$; $p = 0.04$). The KCCQ self-efficacy score was not correlated with NYHA functional class ($r_s = -0.10$; $p = 0.30$) or 6MWT ($r_s = -0.02$; $p = 0.87$). Overall, the correlations of the KCCQ overall summary score with NYHA functional class and 6MWT were slightly stronger as compared with the MLHFQ total score. In contrast to the MLHFQ in which there were no statistically significant correlations with accelerometry, there were modest, statistically significant correlations of the KCCQ clinical summary and physical limitation scores with baseline average accelerometry units.

RESPONSIVENESS TO CLINICAL CHANGE. The changes in MLHFQ and KCCQ scores from baseline to 6 weeks stratified by change in 6MWT distance and NT-proBNP are shown in **Figure 1**. Patients that improved the distance that they could walk from baseline to 6 weeks also had the greatest improvement in HRQOL scores (**Figure 1A**).

On average, patients that experienced improvements in 6MWT distance also had significant improvements in HRQOL scores. Patients that experienced no change in 6MWT still experienced some improvement in HRQOL scores, though of lesser magnitude. Patients that had deterioration in 6MWT distance, on average, had little to no change in HRQOL scores.

Changes in both the MLHFQ total score and KCCQ overall summary score reflected observed changes in 6MWT distance. For patients who experienced improvement, no change, and deterioration in 6MWT distance, the MLHFQ scores improved by an average of 12.83, 7.09, and 1.95 points, respectively. Similarly, KCCQ overall summary scores improved by an average of 10.12, 4.10, and 1.90 points, respectively. Both the MLHFQ physical score and the corresponding KCCQ clinical summary score varied somewhat

TABLE 1 Baseline Patient Characteristics

	Missing	Overall
Age, yrs	0	69.3 ± 9.3
Women	0	63 (57)
White	0	98 (89)
Body mass index, kg/m^2	0	35.6 ± 8.4
Ejection fraction, %	0	63.2 ± 8.5
Comorbidities		
Hypertension	0	99 (90)
Ischemic heart disease	0	68 (62)
Atrial fibrillation	0	39 (35)
Diabetes mellitus	0	43 (39)
COPD	0	16 (15)
Sleep apnea	3	57 (53)
Depression	0	36 (33)
Anemia	1	36 (33)
Medication use		
Loop diuretic	0	72 (65)
ACE inhibitor/ARB	0	70 (64)
Beta-blocker	0	77 (70)
Aldosterone antagonist	0	27 (25)
Anticoagulant or antiplatelet	0	90 (82)
Laboratories		
Creatinine, mg/dl	1	1.2 ± 0.4
NT-proBNP, pg/ml	0	227.2 (103.7–610.1)
Functional status		
NYHA functional class	2 (1 functional class I and 1 functional class IV)	
II		58 (53)
III		50 (45)
6-min walk distance, m	1	311.5 ± 119.1
Baseline HRQOL questionnaires		
MLHFQ*		
Total score	16	44.6 ± 23.7
Physical dimension score	7	22.3 ± 10.2
Emotional dimension score	3	10.0 ± 7.5
KCCQ†		
Overall summary score	0	56.0 ± 24.2
Clinical summary score	0	57.6 ± 22.6
Physical limitation score	0	56.2 ± 24.7
Symptom stability score	0	47.5 ± 17.3
Symptom frequency score	2	58.2 ± 25.1
Symptom burden score	0	59.5 ± 23.9
Total symptom score	0	59.0 ± 23.5
Self-efficacy score	2	77.8 ± 22.7
Quality-of-life score	2	52.8 ± 28.3
Social limitation score	3	56.4 ± 31.7

Values are mean ± SD, n (%), or median (interquartile range). *Minnesota Living with Heart Failure Questionnaire (MLHFQ) scores range from 0 to 105, with lower scores indicative of better health-related quality of life (HRQOL). †Kansas City Cardiomyopathy Questionnaire (KCCQ) scores range from 0 to 100, with higher scores indicative of better health-related quality of life.

ACE = angiotensin-converting enzyme; ARB = angiotensin receptor blocker; COPD = chronic obstructive pulmonary disease; NT-proBNP = N-terminal pro-B-type natriuretic peptide; NYHA = New York Heart Association.

TABLE 2 Internal Consistency of the MLHFQ and KCCQ Domains

HRQOL Instrument	Cronbach's α		
	Baseline	6 Weeks	12 Weeks
MLHFQ			
Physical subscale	0.93	0.94	0.95
Emotional subscale	0.92	0.93	0.95
KCCQ			
Clinical summary	0.91	0.92	0.92
Physical limitation	0.89	0.86	0.87
Total symptom	0.80	0.87	0.86
Self-efficacy	0.72	0.69	0.75
Quality of life	0.87	0.89	0.85
Social limitation	0.92	0.86	0.86

Abbreviations as in Table 1.

according to changes in 6MWT distance, but to a lesser degree than the overall scores.

Unlike the 6MWT, the HRQOL scores did not reflect changes in NT-proBNP, as no consistent pattern of change in HRQOL was observed with change in NT-proBNP (Figure 1B). If HRQOL scores were reflective of changes in NT-proBNP, you would expect the improvement in HRQOL to be greatest in patients where NT-proBNP decreased. However, the magnitude of improvement in HRQOL scores was fairly similar regardless of the change in NT-proBNP. For example, the MLHFQ total score improved by an average of 3.14, 9.76, and 8.83 points in patients who experienced a decrease, no change, and an increase in NT-proBNP, respectively. In comparison, the KCCQ overall summary score improved by 3.90, 6.38, and 4.37 points, respectively.

Responsiveness statistics were subsequently calculated to determine the relative ranking of the QOL measures by improvement or deterioration in 6MWT (Table 4). The MLHFQ total score was the most responsive to improvement in 6MWT, whereas the MLHFQ physical score was most responsive to deterioration in 6MWT. HRQOL domains reflecting emotional aspects of QOL, including the MLFHQ emotional score and KCCQ QOL score, were the least responsive to both improvement and deterioration in 6MWT. As none of the HRQOL scores were responsive to change in NT-proBNP (Figure 1B), ranking of HRQOL measures for responsiveness to change in NT-proBNP was not performed. The responsiveness of the MLHFQ and KCCQ from baseline to 12 weeks are shown in Online Table 4. Similar to the 6-week findings, the MLHFQ total and physical scores were most responsive to changes in 6MWT.

DISCUSSION

The MLHFQ and the KCCQ, the 2 most widely used tools to measure HRQOL in HF, were both developed in patients with HFrEF (8–10). However, patients with HFpEF now comprise one-half of all patients with HF (5,18,19), and differ from patients with HFrEF in pathophysiology and responsiveness to treatment. As such, it is important that any tools used to measure key patient-reported outcomes such as HRQOL in HF are carefully assessed to ensure they are both valid and responsive to clinical change in patients with HFpEF. In this secondary analysis of data from the NEAT trial, we found the MLHFQ and KCCQ to both be reliable, valid, and responsive to clinical change in HFpEF. The KCCQ was more strongly correlated with baseline functional status parameters, whereas the MLHFQ was more responsive to improvement in 6MWT over time. These subtle differences may

TABLE 3 Mean MLHFQ and KCCQ Scores by Baseline NYHA Functional Class and 6MWT Distance

Questionnaire	NYHA Functional Class				6MWT Distance				NT-proBNP			
	II	III	Spearman	p Value	≥300 m	<300 m	Spearman*	p Value	≥227 (Median)	<227 (Median)	Spearman*	p Value
MLHFQ†												
Total score	38.7 ± 24.7	51.5 ± 20.5	0.30	0.003	40.6 ± 22.7	50.4 ± 24.3	−0.23	0.024	41.4 ± 23.8	47.5 ± 23.5	−0.17	0.11
Physical subscale	18.4 ± 10.4	26.0 ± 7.9	0.35	<0.001	19.2 ± 9.7	26.3 ± 9.3	−0.39	<0.001	20.9 ± 11.0	23.6 ± 9.4	−0.16	0.10
Emotional subscale	8.7 ± 7.6	11.0 ± 6.9	0.15	0.13	9.3 ± 7.3	10.6 ± 7.6	−0.11	0.28	8.8 ± 7.4	11.1 ± 7.6	−0.21	0.027
KCCQ‡												
Overall summary	65.5 ± 22.5	46.7 ± 20.9	−0.37	<0.001	63.6 ± 21.3	46.7 ± 23.8	0.38	<0.001	58.4 ± 25.9	53.6 ± 22.4	0.13	0.19
Clinical summary	66.5 ± 20.2	48.9 ± 20.4	−0.37	<0.001	65.9 ± 18.6	47.4 ± 22.3	0.44	<0.001	59.5 ± 24.4	55.8 ± 20.7	0.11	0.25
Physical limitation	66.1 ± 21.6	46.6 ± 22.8	−0.38	<0.001	66.2 ± 19.0	43.7 ± 25.2	0.47	<0.001	57.5 ± 27.4	55.0 ± 21.9	0.07	0.49
Total symptom	66.9 ± 21.9	51.3 ± 21.7	−0.30	0.001	65.6 ± 21.4	51.1 ± 22.7	0.31	0.001	61.5 ± 23.5	56.5 ± 23.5	0.16	0.10
Self-efficacy	80.4 ± 22.2	75.0 ± 23.2	−0.10	0.30	78.8 ± 22.4	76.9 ± 23.2	−0.02	0.87	77.0 ± 23.8	78.5 ± 21.7	−0.02	0.85
Quality of life	61.2 ± 27.8	44.6 ± 25.2	−0.28	0.003	57.8 ± 26.7	46.8 ± 28.8	0.19	0.040	56.3 ± 28.6	49.2 ± 27.8	0.15	0.11

Values are mean ± SD. *6-min walk test (6MWT) distance and NT-proBNP analyzed as continuous variables in calculating correlations. †MLHFQ scores range from 0 to 105, with lower scores indicative of better HRQOL. ‡KCCQ scores range from 0 to 100, with higher scores indicative of better HRQOL.

Abbreviations as in Table 1.

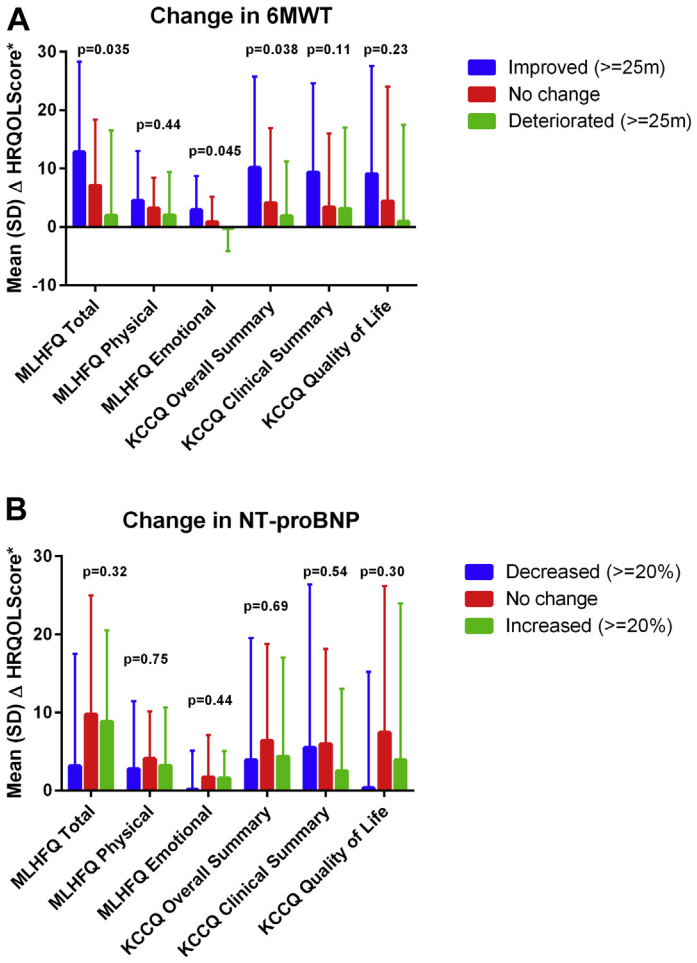
enable preferential selection of a specific tool for a particular clinical or research application.

INTERNAL CONSISTENCY AND RELIABILITY. Whether HRQOL is being assessed in the clinic or as a part of a trial protocol, confidence in the reliability and consistency of results is paramount. In our study, internal consistency reliability as assessed by Cronbach's α was good for the MLHFQ and most KCCQ subscales, with slightly worse reliability observed for the KCCQ self-efficacy subscale. These finding are consistent with prior studies in HFpEF. Rector et al. (20) found the MLHFQ total score and its subscales to have good-to-excellent internal consistency reliability. Similarly, Joseph et al. (6) observed good to excellent internal consistency of most KCCQ subscales with the exception of the self-efficacy subscale, which was less reliable. Overall, these findings confirm that the MLHFQ and the KCCQ are reliable tools to measure HRQOL in patients with HFpEF.

VALIDITY. It is also important to ensure that tools used to measure HRQOL in HFpEF are valid and associated with other conventional clinical criteria. We evaluated the validity of the MLHFQ and KCCQ by comparing them to 3 common criterion standards in HF, NYHA functional class, 6MWT distance, and NT-proBNP, the former 2 of which have been shown to correlate with HRQOL in other populations (21–23). Although both the KCCQ summary and MLHFQ total scores were at least moderately correlated with NYHA functional class and 6MWT distance, the correlations were stronger for the KCCQ. Similarly, only the KCCQ subscales (clinical summary and physical limitation) were significantly correlated with baseline accelerometry. Therefore, it may be preferable to use the KCCQ in situations where HRQOL components such as exercise capacity or functional status are most important, but perhaps cannot be directly measured. Although the total and physical scores for the MLHFQ and KCCQ were correlated with NYHA functional class and 6MWT, the scales most reflective of a patient's perceived emotional well-being (MLHFQ emotional subscale and KCCQ self-efficacy subscale) had weak to no correlation with these measures. This suggests that HRQOL domains indicative of emotional well-being are less reflective of a patient's functional status and should not be used as surrogates for these clinical parameters.

RESPONSIVENESS TO CLINICAL CHANGE. Finally, patient-reported outcome measures such as HRQOL need to be sensitive to clinically meaningful differences in patient outcomes. We observed that both the MLHFQ and KCCQ were responsive to change in 6MWT distance over time. Patients with an improvement in 6MWT distance, on average, experienced large improvements in HRQOL scores. However, even patients with no change in 6MWT distance still experienced a modest improvement in HRQOL,

FIGURE 1 Change in HRQOL Scores by Change in Clinical Parameters



The mean change in health-related quality of life (HRQOL) scores from baseline to 6 weeks is shown stratified by change in **(A)** 6-min walk distance and **(B)** N-terminal pro-B-type natriuretic peptide (NT-proBNP). Cutpoints for improvement (≥25-m increase), no change (<25-m change), and deterioration (≥25-m decrease) in 6MWT distance are based on prior work (2,3). Because prior studies have failed to find a threshold of change in natriuretic peptides that best correlates with changes in HRQOL in heart failure, we defined increased NT-proBNP as an increase of ≥20%, decreased NT-proBNP as a decrease of ≥20%, and no change as change in NT-proBNP <20% (24). Differences in HRQOL scores by change in 6MWT or NT-proBNP categories were examined using analysis of variance (p values shown). *With the Minnesota Living with Heart Failure Questionnaire (MLHFQ), lower scores are indicative of better HRQOL; the opposite is true for the Kansas City Cardiomyopathy Questionnaire (KCCQ). So that changes in MLHFQ and KCCQ scores from baseline to 6 weeks could be easily compared side by side on the same graph, we multiplied each mean change in MLHFQ score by −1. Improved HRQOL from baseline to 6 weeks is now depicted by an increase in score (positive change), whereas worse HRQOL is depicted by a decrease in score (negative change) for both the KCCQ and MLHFQ.

TABLE 4 Responsiveness Statistics and Relative Rankings of HRQOL Measures by Change in 6MWT Distance From Baseline to 6 Weeks

	N	*t* Statistic (Rank)	Effect Size (Rank)	Guyatt's Responsiveness Statistic (Rank)	Standardized Response Mean (Rank)	Median Rank
Improved walk distance (≥+25 m)*						
MLHFQ total	24	0.91 (1)	0.52 (1)	1.13 (1)	0.83 (1)	1.0
MLHFQ physical	32	0.64 (4)	0.43 (3)	0.85 (2)	0.52 (4)	3.5
MLHFQ emotional	30	0.60 (5)	0.38 (5)	0.67 (5)	0.50 (5)	5.0
KCCQ overall	34	0.76 (2)	0.43 (3)	0.79 (3)	0.65 (2)	2.5
KCCQ clinical	34	0.67 (3)	0.44 (2)	0.74 (4)	0.61 (3)	3.0
KCCQ QOL	34	0.49 (6)	0.34 (6)	0.46 (6)	0.49 (6)	6.0
No change in walk distance (+24 to −24 m)*						
MLHFQ total	32	0.50	0.29	0.63	0.63	–
MLHFQ physical	40	0.46	0.32	0.61	0.61	–
MLHFQ emotional	42	0.18	0.11	0.20	0.20	–
KCCQ overall	44	0.31	0.17	0.32	0.32	–
KCCQ clinical	44	0.24	0.14	0.26	0.26	–
KCCQ QOL	42	0.23	0.15	0.22	0.22	–
Deterioration in walk distance (≤−25 m)*						
MLHFQ total	20	0.14 (3)	0.08 (4)	0.17 (3)	0.13 (4)	3.5
MLHFQ physical	23	0.28 (1)	0.19 (1)	0.38 (1)	0.27 (1)	1.0
MLHFQ emotional	27	−0.04 (6)	−0.03 (6)	−0.04 (6)	−0.05 (6)	6.0
KCCQ overall	27	0.14 (3)	0.09 (3)	0.15 (4)	0.20 (3)	3.0
KCCQ clinical	27	0.22 (2)	0.14 (2)	0.24 (2)	0.22 (2)	2.0
KCCQ QOL	27	0.05 (5)	0.03 (5)	0.05 (5)	0.06 (5)	5.0

*Calculated as 6MWT distance at 6 weeks minus 6MWT distance at baseline (m).
Abbreviations as in Tables 1 and 3.

and those who had deterioration in 6MWT, on average, had little to no change in HRQOL. Although no prior studies have examined the responsiveness to clinical change over time of the KCCQ in HFpEF, in patients with HFrEF, KCCQ scores declined when 6MWT distance deteriorated (2). There are no comparable analyses evaluating change in 6MWT distance and MLHFQ scores, though 1 study found that changes in MLHFQ total scores at 6 months reflected changes in patient symptoms and NYHA functional class (20). Our findings suggest that, whereas MLHFQ and KCCQ total and physical scores, in general, reflect changes in 6MWT distance over time, patients who experience a functional deterioration may not report worse HRQOL. Of all scales examined, the MLHFQ total score was the most responsive to improvement in 6MWT distance over time. As such, the MLHFQ total score may be preferred in settings where capturing improvements in exercise capacity and its implications on HRQOL over time are paramount.

Neither the MLHFQ nor the KCCQ correlated with baseline NT-proBNP levels or was responsive to changes in NT-proBNP over time. The lack of responsiveness of HRQOL measures to change in natriuretic peptides aligns with prior work in HFrEF demonstrating that baseline B-type natriuretic peptide and KCCQ are not significantly correlated, and no relationship of changes in B-type natriuretic peptide and KCCQ over time was observed (24). HRQOL has been demonstrated to correlate better than natriuretic peptides with physician- and patient-assessed change in clinical status in HFrEF (3). However, recent work has suggested that less than one-half (48%) of patients with HF rate HF as the primary factor limiting their QOL (25). In particular, patients with HFpEF, who often have more comorbid conditions than those with HFrEF do (26,27), are more likely to rate other medical or nonmedical factors as equally or more important limitations to their HRQOL as their HF (25). Our findings, in the context of prior work, suggest that NT-proBNP, an HF-specific biomarker, may fail to capture the complex factors that affect HRQOL as well as assessments of exercise capacity such as the 6MWT in patients with HFpEF.

STUDY LIMITATIONS AND STRENGTHS. Several potential limitations of this study should be discussed. There was a modest number of patients in the NEAT trial, which limited our ability to identify subtle variations in the performance of the KCCQ and MLHFQ in

subpopulations of HFpEF patients that may exist. Because patients were followed in the NEAT trial for a total of 12 weeks, we could not assess the responsiveness of the KCCQ and MLHFQ to clinical changes over longer time periods. The NEAT trial primarily included patients with NYHA functional class II and III symptoms, and results observed would apply to similar patients. Replication of these findings in larger populations of patients with HFpEF with longer follow-up would be an important next step. Although the KCCQ and MLHFQ are the 2 most commonly used disease-specific HRQOL measures for patients with HF, other general HRQOL questionnaires exist that were not included in this study. However, prior work has suggested that general HRQOL measures do not capture changes in clinical status over time as well as disease-specific measures in HF (3). Although we assessed clinical status using several parameters, including NYHA functional class, 6MWT distance, accelerometry, and NT-proBNP, other measures of HF disease status such as peak oxygen consumption were not assessed. Finally, the NEAT trial excluded patients with selected severe comorbid conditions, including those with critically reduced renal function (estimated glomerular filtration rate <20 ml/min/1.73 m^2), severe anemia, active collagen vascular disease, and those with a non-HF terminal illness. Although enrolled participants still had a significant comorbidity burden (**Table 1**), the NEAT population may not fully reflect the breadth of HFpEF patients seen in clinical practice. However, there are several important strengths. This is the first study to comprehensively evaluate and compare the internal consistency and validity of the MLHFQ and KCCQ in patients with HFpEF. The presence of serial assessments of both HRQOL tools and other clinical parameters enabled us to appraise each tool's responsiveness to change over time.

CONCLUSIONS

Findings from this study suggest that both the MLHFQ and KCCQ are reliable and valid instruments for measurement of HRQOL in patients with HFpEF. In addition, these tools are responsive to changes in some clinical parameters. The KCCQ was more strongly correlated with baseline NYHA functional class and 6MWT distance, whereas the MLHFQ was more responsive to changes in 6MWT distance over time. These findings may be helpful when selecting an appropriate HRQOL tool to use in a particular clinical or research setting. Further studies focused on expanding the use of HRQOL tools in the clinical setting are needed.

ADDRESS FOR CORRESPONDENCE: Dr. Shannon M. Dunlay, Mayo Clinic, 200 First Street Southwest, Rochester, Minnesota 55905. E-mail: dunlay.shannon@mayo.edu.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: These data suggest that the MLHFQ and KCCQ are both valid and reliable instruments to assess HRQOL in patients with HFpEF. Subtle differences in their performance may be helpful to investigators or clinicians in selecting which tool to use in a particular clinical setting or research study.

TRANSLATIONAL OUTLOOK: Future work in larger populations of patients with HFpEF focused on identifying if heterogeneity in performance exists based on patient characteristics is needed. Studies of real-world use of HRQOL tools in clinical practice in HFpEF are needed.

REFERENCES

1. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart Disease and Stroke Statistics-2017 Update: a report from the American Heart Association. Circulation 2017;135:e146–603.

2. Eurich DT, Johnson JA, Reid KJ, Spertus JA. Assessing responsiveness of generic and specific health related quality of life measures in heart failure. Health Qual Life Outcomes 2006;4:89.

3. Spertus J, Peterson E, Conard MW, et al. Monitoring clinical changes in patients with heart failure: a comparison of methods. Am Heart J 2005;150:707–15.

4. Garin O, Herdman M, Vilagut G, et al. Assessing health-related quality of life in patients with heart failure: a systematic, standardized comparison of available measures. Heart Fail Rev 2014;19:359–67.

5. Gerber Y, Weston SA, Redfield MM, et al. A contemporary appraisal of the heart failure epidemic in Olmsted County, Minnesota, 2000 to 2010. JAMA Intern Med 2015;175:996–1004.

6. Joseph SM, Novak E, Arnold SV, et al. Comparable performance of the Kansas City Cardiomyopathy Questionnaire in patients with heart failure with preserved and reduced ejection fraction. Circ Heart Fail 2013;6:1139–46.

7. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2013;62:e147–239.

8. 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;35:1245–55.

9. Rector TS, Cohn JN. Assessment of patient outcome with the Minnesota Living with Heart Failure questionnaire: reliability and validity during a randomized, double-blind, placebo-controlled trial of pimobendan. Pimobendan Multicenter Research Group. Am Heart J 1992;124:1017–25.

10. 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;71:1106–7.

11. Borlaug BA, Redfield MM. Diastolic and systolic heart failure are distinct phenotypes within the heart failure spectrum. Circulation 2011;123:2006–13; discussion 2014.

12. Zakeri R, Levine JA, Koepp GA, et al. Nitrate's effect on activity tolerance in heart failure with preserved ejection fraction trial: rationale and design. Circ Heart Fail 2015;8:221–8.

13. Redfield MM, Anstrom KJ, Levine JA, et al. Isosorbide mononitrate in heart failure with preserved ejection fraction. N Engl J Med 2015;373:2314–24.

14. Rector TS, Tschumperlin LK, Kubo SH, et al. Use of the Living With Heart Failure questionnaire to ascertain patients' perspectives on improvement in quality of life versus risk of drug-induced death. J Card Fail 1995;1:201–6.

15. Kosiborod M, Soto GE, Jones PG, et al. Identifying heart failure patients at high risk for near-term cardiovascular events with serial health status assessments. Circulation 2007;115:1975–81.

16. Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD, Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? Value Health 2007;10 Suppl 2:S94–105.

17. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: Erlbaum; 1988.

18. Redfield MM. Heart failure with preserved ejection fraction. N Engl J Med 2016;375:1868–77.

19. Ho JE, Enserro D, Brouwers FP, et al. Predicting heart failure with preserved and reduced ejection fraction: the International Collaboration on Heart Failure Subtypes. Circ Heart Fail 2016;9.

20. Rector TS, Carson PE, Anand IS, et al. Assessment of long-term effects of irbesartan on heart failure with preserved ejection fraction as measured by the Minnesota Living with Heart Failure questionnaire in the irbesartan in heart failure with preserved systolic function (I-PRESERVE) trial. Circ Heart Fail 2012;5:217–25.

21. Holland R, Rechel B, Stepien K, Harvey I, Brooksby I. Patients' self-assessed functional status in heart failure by New York Heart Association class: a prognostic predictor of hospitalizations, quality of life and death. J Card Fail 2010;16:150–6.

22. Juenger J, Schellberg D, Kraemer S, et al. Health related quality of life in patients with congestive heart failure: comparison with other chronic diseases and relation to functional variables. Heart 2002;87:235–41.

23. Opasich C, Pinna GD, Mazza A, et al. Six-minute walking performance in patients with moderate-to-severe heart failure; is it a useful indicator in clinical practice? Eur Heart J 2001;22:488–96.

24. Luther SA, McCullough PA, Havranek EP, et al. The relationship between B-type natriuretic peptide and health status in patients with heart failure. J Card Fail 2005;11:414–21.

25. Joyce E, Chung C, Badloe S, et al. Variable contribution of heart failure to quality of life in ambulatory heart failure with reduced, better, or preserved ejection fraction. J Am Coll Cardiol HF 2016;4:184–93.

26. Chamberlain AM, St Sauver JL, Gerber Y, et al. Multimorbidity in heart failure: a community perspective. Am J Med 2015;128:38–45.

27. Triposkiadis F, Giamouzis G, Parissis J, et al. Reframing the association and significance of co-morbidities in heart failure. Eur J Heart Fail 2016;18:744–58.

KEY WORDS heart failure, quality of life

APPENDIX For supplemental tables, please see the online version of this paper.