ORIGINAL ARTICLES

Validation of the Spanish Version of the Minnesota Living With Heart Failure Questionnaire

Olatz Garin,[a] Núria Soriano,[b] Aida Ribera,[b] Montse Ferrer,[a,c] Àngels Pont,[a] Jordi Alonso,[a,d] and Gaietà Permanyer[b,c]; on behalf of the HF-QoL Group

[a]Instituto Municipal de Investigación Médica (IMIM)-Hospital del Mar, Barcelona, Spain
[b]Hospital Universitari Vall d'Hebron, Barcelona, Spain
[c]Universitat Autònoma de Barcelona, Barcelona, Spain
[d]Universitat Pompeu Fabra, Barcelona, Spain

Introduction and objectives. The Minnesota Living with Heart Failure Questionnaire (MLHFQ) is the most commonly used instrument for evaluating quality of life in patients with heart failure. It comprises 21 items and 2 dimensions: the physical and the emotional. The aim of this study was to assess the psychometric properties of the Spanish version of the MLHFQ.

Methods. The MLHFQ and the 36-item short form (SF-36) questionnaire were administered 1 and 2 months after discharge to 677 patients who had been hospitalized for heart failure. Patients were classified as either stable (n=245) or unstable (n=103) on the basis of New York Heart Association (NYHA) functional class and 3 other functional capacity variables. Reliability was evaluated using measures of internal consistency (Cronbach's alpha) and reproducibility (the intraclass correlation coefficient [ICC]). Validity was assessed by looking at the scores by NYHA class, and at correlations between scores on MLHFQ and SF-36 dimensions. Responsiveness to change was evaluated using the effect size.

Results. Cronbach's alpha was ≥0.8 for the 3 MLHFQ scores, and the ICC was also large (0.74-0.83). In addition, MLHFQ scores varied significantly with functional class (*P*<.001), and there were intermediate-to-high correlations with the assumed corresponding SF-36 dimensions (0.74-0.52). The observed effect sizes were small or intermediate (0.09-0.44).

Conclusions. The Spanish version of the MLHFQ demonstrated adequate metric properties, comparable to the original. These results support the use of the MLHFQ in Spanish heart failure patients, although it would be advisable to re-evaluate its responsiveness to change.

Key words: *Quality of life. Heart failure. Metric properties. Patient-reported outcome.*

SEE EDITORIAL ON PAGES 233-5

Validación de la versión española del Minnesota Living with Heart Failure Questionnaire

Introducción y objetivos. El Minnesota Living with Heart Failure Questionnaire (MLHFQ) es el instrumento más utilizado para la evaluación de la calidad de vida en pacientes con insuficiencia cardiaca. Contiene 21 ítems y dos dimensiones: física y emocional. El objetivo de este estudio es evaluar las propiedades métricas de la versión española del MLHFQ.

Métodos. Se aplicó, 1 y 2 meses después del alta, el MLHFQ y el SF-36 a 677 pacientes ingresados por insuficiencia cardiaca. A partir de la clasificación NYHA y otras 3 variables de capacidad funcional, se definió a los pacientes como estables (n = 245) o con cambio (n = 103). De la fiabilidad, se estudió: la consistencia interna (alfa de Cronbach) y la reproducibilidad (coeficiente de correlación intraclase [CCI]). La validez se estudió con las puntuaciones según la clase funcional y las correlaciones con las dimensiones del SF-36. La sensibilidad al cambio se evaluó por el tamaño del efecto.

Resultados. El alfa de Cronbach fue ≥ 0,8 en las 3 puntuaciones, y el CCI también fue elevado (0,74-0,83). Las puntuaciones del MLHFQ mostraron diferencias según la clase funcional (p < 0,001), así como correlaciones moderadas-altas con las dimensiones del SF-36 planteadas a priori (0,74-0,52). Los cambios observados fueron pequeños o moderados (0,09-0,44).

Conclusiones. La versión española del MLHFQ ha mostrado unas adecuadas propiedades métricas, igual que la original. Estos resultados respaldan el uso del MLHFQ en pacientes españoles con insuficiencia cardiaca, aunque sería recomendable reevaluar su sensibilidad al cambio.

Palabras clave: *Calidad de vida. Insuficiencia cardiaca. Propiedades métricas. Resultados percibidos por los pacientes.*

ABBREVIATIONS

ES: effect size
HF: heart failure
HRQL: health-related quality of life
ICC: intraclass correlation coefficient
MLHFQ: Minnesota Living with Heart Failure
   Questionnaire
NYHA: New York Heart Association
SD: standard deviation
SF-36: short form, 36 items

INTRODUCTION

The prevalence of heart failure has increased in recent decades.[1] Data from the Framingham study indicate an incidence of heart failure in the population aged over 45 years of 7.2 and 4.7/1000 person/years for males and females, respectively.[2] In developed countries, heart failure is the most frequent cause of hospitalization in patients aged 65 years or over and is the cause of at least 5% of all hospitalizations and 4% of all deaths.[3]

Heart failure has a considerable impact on patients' daily activities, an impact which is comparable to or even greater than that of other chronic diseases such as diabetes or arthritis.[4] The impact of the disease has traditionally been measured using clinical tools, such as the New York Heart Association (NYHA) functional classitification[5] or the 6 minute walk test (6MWT).[6] Health-related quality of life (HRQL) instruments provide means of exploring patient perceptions of the ways in which heart failure affects their daily lives and well-being. Given that clinical indices of severity correlate only weakly or moderately with patient perceptions, HRQL assessments provide additional information which cannot be directly extrapolated from clinical measures.[7]

As heart failure treatments are primarily symptomatic, disease-specific questionnaires for use in these patients have become increasingly important in recent decades.[8] To date, information has been published on the development and validation of 5 questionnaires specifically for use in heart failure, ie, the Minnesota Living with Heart Failure Questionnaire (MLHFQ),[9] the Quality of Life Questionnaire for Severe Heart Failure,[10] the Chronic Heart Failure Questionnaire,[11] the Kansas City Cardiomyopathy Questionnaire,[12] and the Left Ventricular Dysfunction Questionnaire.[13] The most widely known and used of these is the MLHFQ, which has been adapted for use in over 32 languages and has demonstrated good psychometric properties in numerous studies.[14-16]

The MLHFQ was linguistically adapted for use in Spain in 1997 and has been widely used in several

settings.[17-20] Nevertheless, we are not aware of any published studies which have examined the metric properties of the adapted version. The objective of the present study was to assess the feasibility, reliability, validity, and sensitivity to change of the Spanish version of the MLHFQ in daily clinical practice in cardiology outpatient clinics.

METHODS

Study Design

This was a prospective study in which patients admitted for heart failure were recruited consecutively in 50 Spanish hospitals. Patients were followed up for a period of 3 months after discharge in cardiology outpatient clinics.

Patients were considered eligible to participate in the study if they were admitted to hospital for suspected heart failure in a coronary, cardiology, internal medicine, or intensive care unit, and if heart failure was confirmed at discharge as the primary or secondary diagnosis. Inclusion criteria were those of the European Society of Cardiology (symptoms of heart failure and evidence of cardiac dysfunction based on findings from complementary explorations).[21] Exclusion criteria were: *a)* heart failure secondary to a reversible acute cause (supraventricular tachyarrhythmia which reverts to a sinus rhythm, hyperthyroidism); *b)* heart failure or acute pulmonary oedema secondary to serious valvulopathy requiring surgery; *c)* presence of serious concomitant illness (chronic kidney disease requiring renal replacement therapy, in treatment for a neoplasm) or a diagnosis of cor pulmonale; and *d)* unable to participate due to clinical status.

The study was approved by the Ethics Committee of the Vall d'Hebron Hospital in Barcelona.

On hospital admission, demographic and clinical data (illness history and co-morbidity, severity, and etiology of heart failure, functional capacity) were collected, explorations were performed, and treatment prescribed. Functional capacity was described using the NYHA classification together with 3 questions using a dichotomous yes/no response (Do you regularly walk outside? Do you perform any recreational activity requiring physical exertion? Do you refrain from exerting yourself?).

HRQL data was collected and a clinical evaluation (rehospitalizations, visits, and diagnostic tests, functional status, and changes in treatment) at baseline, which was 1 month after discharge. The same data were collected at a second visit 2 months after the first.

Quality of Life Questionnaires

The MLHFQ was developed in the USA by T. Rector.[12] It is a self-administered questionnaire consisting of 21 items. It provides an overall score as well as a score for the 2 dimensions of physical (8 items) and emotional (5

items) health. Response options are from 0 (no impact on HRQL) to 5 (maximum impact on HRQL). Overall (0-105) and dimension (physical, 0-40; emotional, 0-25) scores are obtained by summing responses to each of the items. Scores can be imputed as long as there are fewer than 4 missing values on the physical dimension, fewer than 3 on the emotional dimension, and fewer than 11 for the overall score.

The generic SF-36 questionnaire was administered alongside the disease-specific MLHFQ. The SF-36 health questionnaire can be administered in the general population and different patient groups,[22] and has been used to evaluate several interventions in heart failure.[17,19] It includes 36 questions which measure the following 8 dimensions of health: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health.[23] A score is obtained for each dimension ranging from 0 (worst) to 100 (best health). The instrument also generates 2 summary scores representing mental and physical health. These are standardized to a mean of 50 and a standard deviation [SD] of 10 using Spanish general population reference scores.[24] Summary scores above 50 indicate better HRQL than the general population; those below 50 indicate poorer HRQL.

Sub-Groups

Patients were divided into 2 sub-groups based on the degree of change on the 4 functional capacity variables (NYHA and the 3 additional questions) between the 2 visits. Patients who did not show any change on any of the 4 variables were considered stable and were used to evaluate test-retest reliability. Patients who improved or deteriorated on at least 2 of the 4 variables were included in the sub-group used to analyze sensitivity to change.

Statistical Analysis

Sub-group socio-demographic and clinical characteristics, and HRQL scores, were compared using parametric or non-parametric tests depending on the distribution of the continuous variables. The χ^2 test was used to compare sub-groups on categorical variables.

The observed range of scores on the 2 HRQL questionnaires was calculated for the baseline visit. Feasibility was assessed by calculating the percentage of patients per dimension with at least 1 missing value. Floor and ceiling effects (the proportion of patients with the maximum and minimum score, respectively) were obtained for each score. Reliability was assessed by calculating: a) internal consistency (estimated using Cronbach's alpha coefficient,[25]); and b) test-retest reliability (estimated using the intraclass correlation coefficient [ICC][26]). Cronbach's alpha measures the degree of homogeneity among items in a dimension at a single administration. In the present study, it was calculated using data for the whole sample from the baseline evaluation. The ICC is a measure of agreement and was calculated using data from the 2 assessments in the reproducibility sub-sample. Both Cronbach's alpha and the ICC take values between 0 and 1. A value of .7 has been suggested as the threshold for comparisons at group level, while for individual level comparisons a value of α=.9 is considered appropriate.[27]

Construct validity refers to the extent to which scores correlate in expected ways with other clinical or HRQL measures.[27] In order to examine the pattern of HRQL scores across known groups defined by different levels of clinical severity, figures were constructed showing the means and 95% confidence intervals (CI) for each NYHA functional class. The construct validity of the MLHFQ was assessed using a Spearman correlation matrix of the SF-36 and MLHFQ dimensions (multi-trait multi-method matrix).[28] Expected correlations between the 2 HRQL instruments were categorized as convergent or discriminant. Convergent validity refers to the idea that correlations between different instruments measuring similar concepts should be moderate to high, ie, >0.4 and >0.6, respectively. We hypothesized that the highest correlations would be observed between: a) the physical dimension of the MLHFQ and the SF-36 physical functioning, role physical, and physical summary scores; and b) the emotional dimension of the MLHFQ and the role emotional, mental health, and mental component summary score of the SF-36. To show discriminant validity, there should be low correlations between instruments which aim to measure different traits. We therefore hypothesized a priori that there would be low correlations between: the SF-36 physical functioning, role physical, and physical summary scores, and the emotional dimension of the MLHFQ as well as between the physical dimension of the MLHFQ and the SF-36 role emotional, mental health, and mental summary scores.

Analysis of the MLHFQ's sensitivity to change was performed using data from the sub-groups who reported improvement and deterioration. First, mean scores from the visits 1 and 3 months after discharge were compared using Wilcoxon t test. The effect size (ES) was calculated for both the MLHFQ and the SF-36 based on the change in score between the 2 visits.[29] The effect size (ES) is equivalent to mean change / baseline SD. An ES of >0.8 is considered high; one of 0.5 moderate, and one close to 0.2 is considered low.

RESULTS

A total of 677 patients with heart failure were included in the study. The final sample had a mean age (SD) of 69.6 (11.9) years; 61% were male. Patients were generally classified in NYHA groups I, II, and III (19.6%, 53.1%, and 25.2%, respectively). All HRQL instruments were
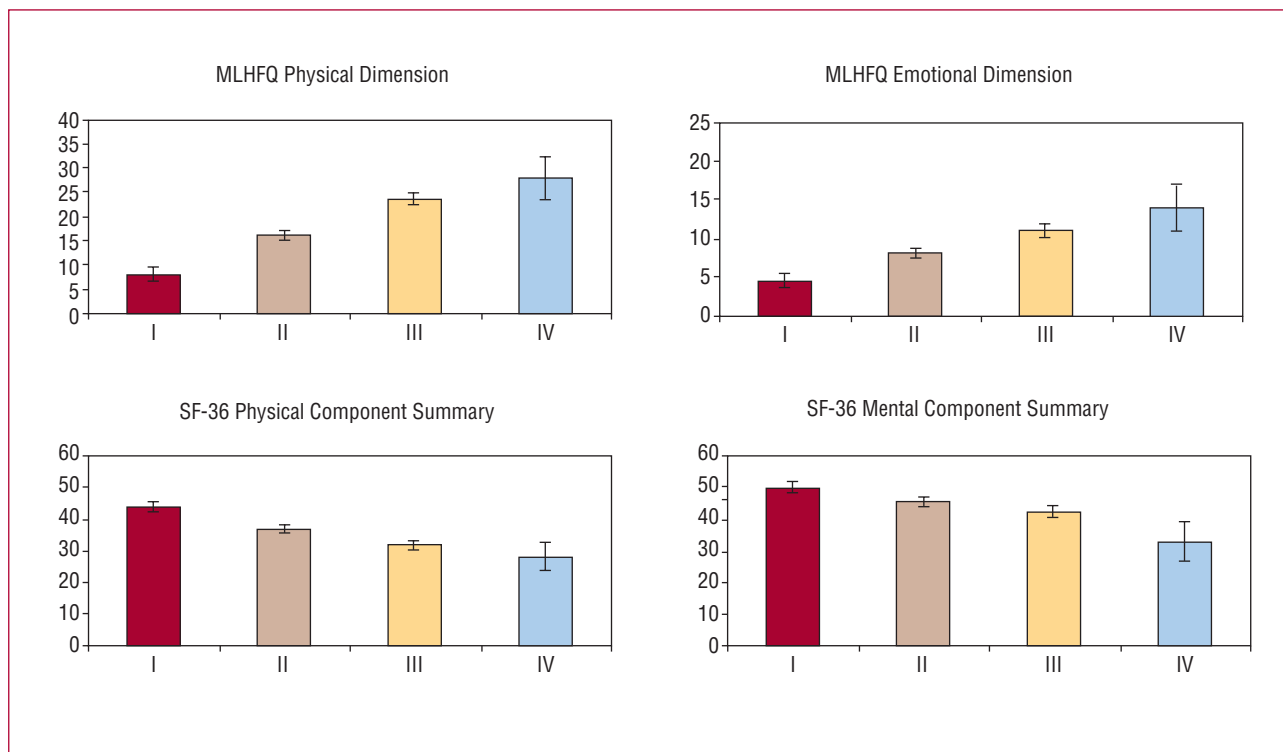
Figure 1. Scoring gradient for the MLHFQ physical and emotional domains and SF-36 summary scores by NYHA classification. Baseline assessment.

completed by patients at the study visit and were only administered by health care personnel when required.

Table 1 shows the sample's socio-demographic and clinical characteristics, as well as HRQL scores for the overall sample and for the reproducibility (n=245) sub-group and for the sub-groups of patients who improved (n=60) and deteriorated (n=43). There were statistically significant differences between sub-groups on the following variables: distribution on the NYHA, age, SF-36 mental summary score, on 3 dimensions of the SF-36, and on the MLHFQ scores.

Observed scores on the MLHFQ and SF-36 covered the full theoretical range (Table 2). The rate of missing responses was practically zero in the 2 MLHFQ dimensions, and was only high for the overall score (22.5%). MLHFQ ceiling and floor effects were very low, though there were substantial ceiling effects on 4 of the SF-36 dimensions. Cronbach's alpha was high for all dimensions, ranging from .817-.915 and .70-.93 on the MLHFQ and SF-36, respectively. The ICC was >0.7 for the 3 MLHFQ scores and close to 0.6 in the majority of SF-36 dimensions.

The difference in score between the different NYHA classes was statistically significant (Kruskal-Wallis, $P<.001$) in all cases (Figure 1), with mean scores for the MLHFQ physical dimension ranging from 7.9 (8.4) for class I to 27.8 (8.3) for class IV. Mean SF-36 physical summary scores were 43.7 (8.6) and 28.1 (8.4) in functional classes I and IV, respectively. When

discriminant validity was assessed using scores from the second evaluation, the results were very similar ($P<.001$) (Figure 2).

The correlation matrix between the dimensions of the SF-36 and the MLHFQ shows that all of the correlations which had previously been hypothesized to be moderate or high (Table 3) were >0.52, with the exception of the correlation between mental health and the emotional domain of the MLHFQ, which was 0.39. Correlations which were expected a priori to be low (discriminant validity) were <0.5 (Table 3).

Between the 2 study visits, approximately 20% of patients remained stable and change in NYHA functional class was similar in the improvement and deterioration sub-groups: the majority changed by only 1 class (70% of those who improved and 69.8% of those who deteriorated) and none of the patients changed by more than 2 classes. In the sub-group of patients who improved, ES were quite low (Table 4), the largest being seen on the physical domain and the MLHFQ overall score (0.42 and 0.41, respectively). In the sub-group of patients who worsened, ES were even smaller, from –0.09 to –0.26 on the MLHFQ.

DISCUSSION

The Spanish version of the MLHFQ has demonstrated adequate measurement properties which are similar to those of the original version. The excellent results in

TABLE 1. Sociodemographic and Clinical Characteristics and Baseline Health-Related Quality of Life Scores

	Patients Included (n=677)	Reproducibility Sub-sample (n=245)	Improvement Sub-sample (n=60)	Deterioration Sub-sample (n=43)
Age, mean (SD), median	69.6 (11.9), 71.2	68.2 (12.1), 70.4	72.5 (12.4), 75	68.2 (11.4), 70.3
Males, n (%)	413 (61)	158 (64.5)	38 (63.3)	26 (60.5)
NYHA, n (%)				
I	127 (19.6)	57 (23.3)	1 (1.7)	13 (30.2)
II	344 (53.1)	135 (55.1)	32 (53.3)	29 (67.4)
III	163 (25.2)	52 (21.2)	25 (41.7)	1 (2.3)
IV	14 (2.2)	1 (0.4)	2 (3.3)	0
Prior diagnosis of heart failure, n (%)	407 (60.7)	146 (59.6)	41 (69.5)	23 (54.8)
Smoking, n (%)				
Non-smoker	328 (48.4)	116 (47.3)	30 (50)	21 (50)
Ex-smoker	246 (36.6)	93 (38)	21 (35)	15 (35.7)
Current smoker	98 (14.6)	36 (14.7)	9 (15)	6 (14.3)
Cause of admission, n (%)				
Heart failure	575 (85.7)	210 (85.7)	52 (88.1)	33 (80.5)
Acute myocardial infarct	57 (8.5)	22 (9)	3 (5.1)	4 (9.8)
Others	39 (5.8)	13 (5.3)	4 (6.8)	4 (9.8)
Ejection fraction, n (%)				
≥ 50%	213 (33.3)	79 (33.1)	12 (20.7)	10 (24.4)
40-49%	118 (18.4)	42 (17.6)	14 (24.1)	9 (22)
30-39%	152 (23.8)	65 (27.2)	14 (24.1)	7 (17.1)
<30%	157 (24.5)	53 (22.2)	18 (31)	15 (36.6)
Etiology at discharge, n (%)				
Hypertensive	279 (43.5)	99 (41.8)	30 (54.5)	18 (45)
Ischemic	294 (46)	106 (44.9)	22 (40)	17 (41.5)
Valvular	125 (19.5)	40 (16.9)	11 (20)	8 (19.5)
Cardiomyopathy	171 (26.8)	73 (30.9)	17 (30.9)	12 (28.6)
Others	59 (9.2)	22 (9.2)	3 (5.5)	4 (10)
History of comorbidity, n (%)				
High blood pressure	453 (67.3)	167 (68.2)	41 (68.3)	28 (66.7)
Diabetes mellitus	293 (43.5)	101 (41.2)	24 (40)	19 (45.2)
COPD	147 (21.8)	50 (20.4)	12 (20)	11 (25.6)
Coronary heart disease	276 (41)	95 (38.8)	22 (36.7)	16 (38.1)
Valve disease	158 (23.5)	54 (22)	13 (22)	10 (23.3)
Cardiomyopathy	168 (25)	76 (31)	17 (29.3)	12 (27.9)
CA×AF	284 (42.1)	102 (41.6)	27 (45)	13 (31)
Other diseases	233 (34.5)	81 (33.1)	23 (38.3)	16 (37.2)
SF-36, mean (SD), median				
Physical functioning	49.2 (28.3), 50	54.6 (29.1), 55	47.8 (27.1), 50	54.7 (24), 55
Role physical	35.7 (41.5), 0	36.3 (42), 25	36.3 (42), 12.5	46.5 (41), 50
Role emotional	67.6 (43.8), 100	69.5 (43), 100	58.3 (45.8), 83.3	66.3 (44.3), 100
General health	41.8 (20.1), 40	43.7 (19.9), 42	41.1 (20.3), 42	39.5 (21.3), 40
Vitality	47.1 (23.9), 50	50.6 (23.8), 50	45.4 (23.2), 50	45.8 (21.2), 50
Social functioning*	66.8 (29.3), 75	71.1 (27.2), 75	55.5 (32.9), 62.5	66.0 (28.1), 62.5
Mental health*	59.3 (21.2), 60	62.7 (20.7), 64	55.4 (21.9), 52	55.2 (20.2), 56
Bodily pain*	65.8 (30.3), 62	71.2 (29), 74	60.3 (30.3), 56	68.8 (27), 62
PCS	36.7 (10), 36.7	38.2 (10.1), 38.4	36.4 (9.8), 34	39.4 (9.4), 40.1
MCS*	44.7 (12.5), 46.3	45.9 (12.2), 48.8	40.8 (12.6), 41.3	42.1 (13.3), 43
MLHFQ, mean (SD), median				
Physical*	16.9 (10.2), 18	15.0 (10.6), 15	18.6 (9.3), 20	15.0 (8.6), 17
Emotional*	8.4 (6), 8	7.5 (6.3), 6	9.6 (6.5), 10.5	8.2 (5.4), 9
Total*	37.1 (20.9), 39	32.2 (21.9), 31	43.7 (19.6), 47.5	33.8 (19.3), 29

CA×AF indicates complete arrhythmia due to atrial fibrillation;√MCS, mental component summary;≈MLHFQ, Minnesota Living with Heart Failure Questionnaire; NYHA, New York Heart Association; PCS, physical component summary.
*Statistically significant differences (using Kruskal-Wallis or χ²) between sub-samples.
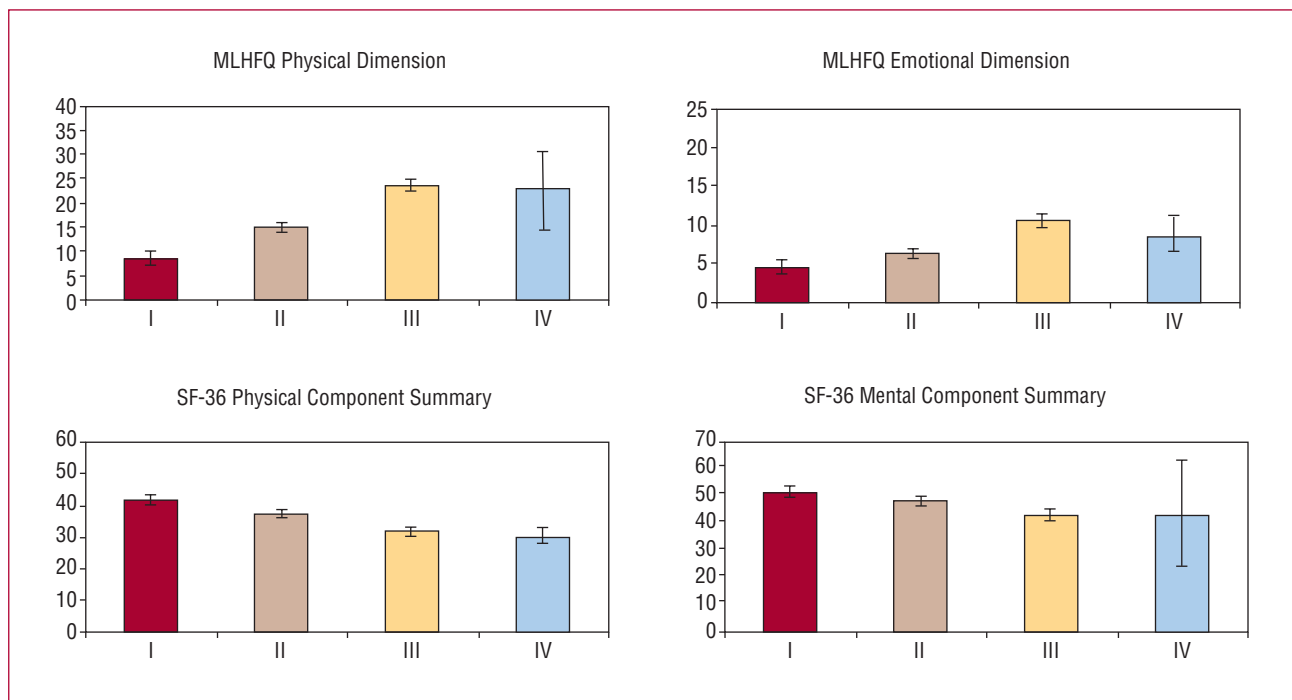
Figure 2. Scoring gradient for the MLHFQ physical and emotional domains and SF-36 summary scores by NYHA classification. Second assessment.

TABLE 2. **Score distributions and Reliability Coefficients for the MLHFQ and SF-36 for the Total Sample (n=677)**

Dimensions	Missing Values, %	Observed Range	Floor Effect, %	Ceiling Effect, %	Cronbach's (n=653)	ICC*
MLHFQ						
Physical	3.5	0-40	0.6	4.6	.90	0.79
Emotional	1.2	0-25	0.6	8.5	.82	0.74
Total	22.5	0-105	0.2	0.6	.91	0.83
SF-36						
Physical functioning	3.8	0-100	4	1.6	.92	0.82
Role physical	2.1	0-100	50.4	21.8	.89	0.62
Role emotional	1.9	0-100	26.5	61.3	.86	0.55
General health	3.0	0-100	1.1	0.2	.70	0.72
Vitality	2.1	0-100	3.3	1.2	.84	0.68
Social functioning	4.3	0-100	3.3	28.3	.78	0.65
Mental health	2.5	4-100	0	2.7	.93	0.70
Bodily pain	0.9	0-100	2.2	32.5	.81	0.57
PCS	–	12.2-60.4	–	–	–	0.73
MCS	–	12.2-68.7	–	–	–	0.65

ICC indicates intraclass correlation coefficient; PCS, physical component summary; MCS, mental component summary; MLHFQ, Minnesota Living with Heart Failure Questionnaire.
*Coefficient calculated using data from the reproducibility sample (n=245).

terms of reliability and validity were particularly noteworthy. The study results support the use of the MLHFQ in Spain as well as its use in international comparisons.

The analysis of the MLHFQ's feasibility showed that there were virtually no missing responses except on 2 items referring to the respondent's profession and sexual activity. These 2 items are not included in any of the questionnaire dimensions, but they do contribute to the overall score. Despite the results on these 2 items, the study indicates that the Spanish version of the MLHFQ is feasible for use in heart failure patients. The distribution

TABLE 3. Spearman Correlation Matriz (Multi-Trait Multi-Method) Between Minnesota Living With Heart Failure Questionnaire (MLHFQ) and SF-36 Dimensions

| | MLHFQ | | |
	Physical	Emotional	Total
MLHFQ			
Physical			
Emotional	0.655		
Total	0.919	0.820	
SF-36			
Physical functioning	−0.738[a]	−0.481[b]	−0.660
Role physical	−0.523[a]	−0.337[b]	−0.457
Bodily pain	−0.377	−0.437	−0.417
General health	−0.422	−0.395	−0.431
Vitality	−0.654	−0.532	−0.651
Social functioning	−0.567	−0.582	−0.624
Role emotional	−0.464[b]	−0.630[a]	−0.586
Mental health	−0.394[b]	−0.390[a]	−0.410
PCS	−0.633[a]	−0.350[b]	−0.547
MCS	−0.412[b]	−0.583[a]	−0.528

MCS indicates mental component summary; PCS, physical component summary;
[a]Correlations expected a priori to be moderate; convergent validity.
[b]Correlations expected a priori to be low; discriminant validity.

of the scores clearly illustrates some of the advantages of disease-specific over generic instruments. The small ceiling and floor effects and the use of the full range of scores in a sample which covers the full range of severity, such as that included in the present study, indicate that the questionnaire addresses problems of relevance to these patients and suggest that the instrument is likely

to detect improvement or deterioration. The high percentage of patients with the maximum possible score (ceiling effect) on several dimensions of the SF-36 is, in part at least, a reflection of the instrument's lack of relevance for patients with this condition. On the other hand, the 2 SF-36 role dimensions have shown high ceiling effects in several populations and is one of the reasons for recent modifications to the response scale used in these dimensions in version 2 of the SF-36.[30-32]

The MLHFQ also showed excellent reliability, both in terms of internal consistency and reproducibility, with reliability coefficients over the minimum recommended standard[27] on all dimensions. On the physical dimension and the overall score, Cronbach's alpha was over .90, which has been proposed as the standard for individual level comparisons. The ICC was also >0.7 for all scores. On the other hand, the 95% CI for the SF-36 mean scores for the 4 NYHA classes overlapped, which contrasts with the generally independent means observed for the MLHFQ. This indicates the MLHFQ's greater ability to discriminate between patients according to the degree of functional impairment, whilst the correlations observed between the 2 instruments provide evidence of the construct validity of the physical and emotional domains of the MLHFQ.

The effect sizes observed on the MLHFQ physical dimension and overall score (both were close to 0.4) can be considered moderate, according to Cohen's criteria,[29,33] though they were larger than those observed on the SF-36 dimensions (0.03-0.33). The greater capacity of the MLHFQ to detect change supports the hypothesis, which has also been confirmed in other studies, that disease-specific instruments are more sensitive to change than generic instruments.

TABLE 4. Estimates of Sensitivity to Change for the MLHFQ and SF-36 in the Improvement, and Deterioration Sub-Groups

| | Improvement (n=60) | | | Deterioration (n=43) | | |
	Change, Mean (SD)	P (Wilcoxon's t)	ES	Change, Mean (SD)	P (Wilcoxon's t)	ES
MLHFQ						
Physical	3.96 (8.97)	.002	0.42	−2.24 (8.62)	.052	−0.26
Emotional	1.27 (6.2)	.130	0.19	−0.69 (6.33)	.674	−0.13
Total	8.22 (19.79)	.010	0.41	−1.89 (20.37)	.342	−0.09
SF-36						
Physical functioning	6.23 (23.89)	.105	0.23	−11.15 (26.37)	.011	−0.48
Role physical	13.36 (42.2)	.019	0.33	−10.26 (46.85)	.201	−0.24
Role emotional	5.75 (50.39)	.408	0.13	−11.54 (51.75)	.211	−0.26
General health	−0.56 (19.43)	.893	−0.03	1.68 (26.81)	.871	0.08
Vitality	4.33 (21.86)	.123	0.18	−2.37 (20.29)	.404	−0.11
Social functioning	9.43 (34.17)	.063	0.28	0.32 (35.64)	.931	0.01
Mental health	3.44 (18.01)	.047	0.16	−1.26 (16.72)	.716	−0.06
Bodily pain	1.45 (31.42)	.545	0.05	−9.84 (36.38)	.110	−0.36
PCS	2.00 (9.93)	.114	0.20	−2.78 (11.82)	.105	−0.31
MCS	2.22 (13.02)	.114	0.18	−0.51 (12.61)	.667	−0.04

ES indicates effect size; MCS, mental component summary; MLHFQ, Minnesota Living with Heart Failure Questionnaire; PCS, physical component summary.

Study Limitations

One of the study limitations was that the analysis of test-retest reliability was based on data collected after a 2 month interval, which is a longer time than is generally recommended in this type of design. Over that length of time, there may have been changes in the treatment or in the patient's situation. Furthermore, assessment of the stability of the patient's condition was based on physician ratings of functioning (among other variables), and not only on patient self-report using the health status transition item. The latter may have been more appropriate given that physician and patient ratings do not correlate strongly.[34,35] These 2 methodological characteristics may have led us to underestimate the reproducibility of the Spanish version of the MLHFQ, which may actually be greater than that observed here. Patient improvement or deterioration was likewise measured indirectly, as it was assumed that a change in functional status would be accompanied by a change in self-perceived HRQL, despite the fact that the metric characteristics of the NYHA classification are not well-known,[36] there is considerable variability in its use, and little evidence regarding its capacity to detect a minimum clinically important difference. When evaluating sensitivity to change, the inclusion of a higher proportion of symptomatic patients could provide a more homogeneous distribution across the four NYHA functional classes. It would also be useful to employ a pre-post design with an intervention which would produce a clear improvement in health status. Despite these limitations, the coefficients obtained showed that the Spanish version of the MLHFQ is sensitive to change, and more so than the SF-36.

The study results cannot be considered to provide a representative description of the HRQL of patients hospitalized for heart failure in Spain. Nevertheless, the variety of patients included provides some support for the external validity of the Spanish version's psychometric characteristics. It can therefore be affirmed that the MLHFQ is appropriate for measuring HRQL in patients with heart failure with a range of characteristics.

CONCLUSIONS

This study has shown that the MLHFQ has excellent reliability and validity and moderate sensitivity to change when used to evaluate HRQL in patients with heart failure. Given its cross-cultural characteristics, it will allow for comparisons between countries and will provide a particularly useful measure of HRQL in multi-center international studies.

ACKNOWLEDGEMENTS

HF-QOL GROUP

J. Ariza, J. Fernández, V. López, R. Calvo, P. Bureo, J. Carretero; A. Bayes, D. Gil, C. Ligero, J. Comín, P. Cabero, J. Roure, G. Peñarrojas, M.A. Paz, S. Castro, J. Roca, L. Per-digón, J.A. Ruiz, D. Jiménez, V. Bertomeu, A. Mateu, A. Carrión, S. Martí, A.M. Rubio, J. García, J. Blanquer, J.C. Vargas, C. Pérez, M.A. García, L. Pérez, C. Borasteros, F. Taboada, A. Grande, A.I. Huelmos, J. Bilbao, A. Melero, A. Díaz, J.L. Diago, A. Navarro, J.F. Sotillo, J. Rovira, J.A. Velasco, A. Chaume, D. Atienza, A. Salvador, P. Baello, J. Muñoz, V. Ruiz, M.J. Fombella, J.M. Cerqueiro, E. Freire, J. Jiménez, C. Hidalgo, F. Santolaria, M. Rodríguez, O. Afonso, I. Lekuona, J.A. Alarcón, A. Pérez, J. Marasa, A. del Río, T.Soriano, E. Roig, I. Vallejo, A. Álvarez, J. Julià, R. Bagà, J. Mesquida, A. Tobaruela, J.M. Lomas, A. Martínez, A. Aguilera, and A.M. Campos.

REFERENCES

1. Muñiz J, Crespo MG, Castro A. Insuficiencia cardiaca en España. Epidemiología e importancia del grado de adecuación a las guías de práctica clínica. Rev Esp Cardiol. 2006;6 Suppl F:F2-8.
2. Ho KK, Pinsky JL, Kannel WB, Levy D. The epidemiology of heart failure: the Framingham Study. J Am Coll Cardiol. 1993;22:A6-13.
3. Kannel WB, Belanger AJ. Epidemiology of heart failure. Am Heart J. 1991;121:951-7.
4. Alonso J, Ferrer M, Gandek B, Ware JE, Aaronson NK, Mosconi P, et al. Health-related quality of life associated with chronic conditions in eight countries. Results from the International Quality of Life Assessment (IQOLA) project. Qual Life Res. 2004;13: 283-98.
5. Rossi G. [Nomenclature and diagnostic criteria in cardiology set by the New York Heart Association. Considerations on the 6th edition]. Cuore Circ. 1967;51:287-93.
6. Demers C, McKelvie RS, Negassa A, Yusuf S. Reliability, validity, and responsiveness of the six-minute walk test in patients with heart failure. Am Heart J. 2001;142:698-703.
7. Alonso J, Ferrer M. Enfermedad pulmonar obstructiva crónica y salud percibida. Med Clin (Barc). 1998;111:580-2.
8. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. Med Care. 1989;27:S217-32.
9. Rector TS, Kubo SH, Cohn JN. Patients' self-assessment of their congestive heart failure. Part 2: content, reliability and validity of a new measure, The Minnesota Living with Heart Failure Questionnaire. Heart Failure. 1987;3:198-209.
10. Wiklund I, Lindvall K, Swedberg K, Zupkis RV. Self-assessment of quality of life in severe heart failure. An instrument for clinical use. Scand J Psychol. 1987;28:220-5.
11. 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;4:101-7.
12. 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.
13. O'Leary CJ, Jones PW. The left ventricular dysfunction questionnaire (LVD-36): reliability, validity, and responsiveness. Heart. 2000;83: 634-40.
14. Middel B, Bouma J, de Jongste M, van Sonderen E, Niemeijer MG, Crijns H, et al. Psychometric properties of the Minnesota Living with Heart Failure Questionnaire (MLHF-Q). Clin Rehabil. 2001;15:489-500.

15. Bennett SJ, Oldridge NB, Eckert GJ, Embree JL, Browning S, Hou N, et al. Comparison of quality of life measures in heart failure. Nurs Res. 2003;52:207-16.
16. Pollock SG, Lystash J, Tedesco C, Craddock G, Smucker ML. Usefulness of bucindolol in congestive heart failure. Am J Cardiol. 1990;66:603-7.
17. Rodriguez-Artalejo F, Guallar-Castillon P, Pascual CR, Otero CM, Montes AO, Garcia AN, et al. Health-related quality of life as a predictor of hospital readmission and death among patients with heart failure. Arch Intern Med. 2005;165:1274-9.
18. Parajón T, Lupón J, González B, Urrutia A, Altimir S, Coll R, et al. Aplicación en España del cuestionario sobre calidad de vida «Minnesota Living With Heart Failure» para la insuficiencia cardíaca. Rev Esp Cardiol. 2004;57:155-60.
19. Ojeda S, Anguita M, Delgado M, Atienza F, Rus C, Granados AL, et al. Short- and long-term results of a programme for the prevention of readmissions and mortality in patients with heart failure: are effects maintained after stopping the programme? Eur J Heart Fail. 2005;7:921-6.
20. Guallar-Castillón P, Magariños-Losada MM, Montoto-Otero C, Tabuenca AI, Rodríguez-Pascual C, Olcoz-Chiva M, et al. Prevalencia de depresión, y factores biomédicos y psicosociales asociados, en ancianos hospitalizados con insuficiencia cardiaca en España. Rev Esp Cardiol. 2006;59:770-8.
21. Remme WJ, Swedberg K. Guidelines for the diagnosis and treatment of chronic heart failure. Eur Heart J. 2001;22:1527-60.
22. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30:473-83.
23. Vilagut G, Ferrer M, Rajmil L, Rebollo P, Permanyer-Miralda G, Quintana JM, et al. El cuestionario de salud SF-36 español: una década de experiencia y nuevos desarrollos. Gac Sanit. 2005;19:135-50.
24. Alonso J, Regidor E, Barrio G, Prieto L, Rodríguez C, de La Fuente De Hoz L. Valores poblacionales de referencia de la versión española del Cuestionario de Salud SF-36. Med Clin (Barc). 1998;111:410-6.
25. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16:297-334.
26. Prieto L, Lamarca R, Casado A. La evaluación de la fiabilidad en las observaciones clínicas: el coeficiente de correlación intraclase. Med Clin (Barc). 1998;110:142-5.
27. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res. 2002;11:193-205.
28. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull. 1959;56:81-105.
29. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale: Lawrence Elrbaum Associates; 1988.
30. Jenkinson C, Stewart-Brown S, Petersen S, Paice C. Assessment of the SF-36 version 2 in the United Kingdom. J Epidemiol Community Health. 1999;53:46-50.
31. Taft C, Karlsson J, Sullivan M. Performance of the Swedish SF-36 version 2.0. Qual Life Res. 2004;13:251-6.
32. Ware JE Jr, Kosinski M, Turner-Bowker DM, Gandek B. How to score version 2 of the SF-12 Health Survey (with a supplement documenting version 1). Lincoln: RI; 2002.
33. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care. 1989;27:S178-89.
34. Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. J Clin Epidemiol. 1992;45:743-60.
35. Detmar SB, Muller MJ, Schornagel JH, Wever LD, Aaronson NK. Health-related quality-of-life assessments and patient-physician communication: a randomized controlled trial. JAMA. 2002;288:3027-34.
36. 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;31:262-70.