Validation of the Minnesota Living With Heart Failure Questionnaire in Primary Care

José C. Naveiro-Rilo*, Dulcinea M. Diez-Juárez*, Agustín Romero Blanco*, Francisco Rebullol-Gutiérrez*, Ana Rodríguez-Martínez* and Miguel A. Rodríguez-García*

*Gerencia de Atención Primaria de León, León, Spain
Servicio de Cardiología, Hospital de León, León, Spain

Introduction and objectives. To evaluate the applicability, internal consistency and validity of the Minnesota Living with Heart Failure Questionnaire (MLHFQ) when used in primary care, compared with the Short Form-36 (SF-36) health survey.

Methods. The two questionnaires were administered to 589 patients with chronic heart failure who were registered with 97 primary care physicians. The applicability, internal consistency and validity of the MLHFQ were evaluated and comparisons were made with the SF-36.

Results. More than 90% of patients completed the questionnaires. The percentage of uncompleted items was low. Cronbach’s alpha ranged from 0.79 to 0.94 for the various MLHFQ dimensions. Exploratory factorial analysis identified two factors that explained 65.8% of the variance. Moderate to good correlations were observed between similar dimensions of the MLHFQ and SF-36 (correlation coefficient −0.43 to −0.73). There were significant associations between scores on the MLHFQ and clinical measures of disease severity.

Conclusions. When used in primary care, the MLHFQ had a high level of acceptability and good psychometric properties compared with the SF-36. Consequently, it would be useful for assessing health-related quality of life in patients with chronic heart failure.

Key words: Chronic heart failure. Quality of life. Validation. Minnesota Living with Heart Failure Questionnaire.

Validation del Minnesota Living with Heart Failure Questionnaire en atención primaria

Introducción y objetivos. Evaluar la aplicabilidad, la consistencia interna y la validez del Minnesota Living with Heart Failure Questionnaire (MLHFQ) en atención primaria, comparándolo con el Short-Form Health Survey (SF-36).

Métodos. Se aplicaron ambos cuestionarios a 589 pacientes con insuficiencia cardiaca crónica documentada atendidos por médicos de atención primaria. Analizamos la factibilidad, la consistencia interna y la validez del MLHFQ comparado con el SF-36.

Resultados. Respondió los cuestionarios más del 90% de la muestra. El porcentaje de ítems no respondido es bajo. El coeficiente alfa de Cronbach oscila entre 0,79 y 0,94 para las dimensiones del MLHFQ. Del análisis factorial exploratorio, se extraen dos factores que explican una varianza total del 65.8%. Los coeficientes de correlación entre dimensiones similares del MLHFQ y el SF-36 fueron de moderados a altos (−0,43 a −0,73). Las puntuaciones del MLHFQ se asocian significativamente con variables clínicas de gravedad.

Conclusiones. En atención primaria el MLHFQ, comparado con el SF-36, muestra buena aceptabilidad y buenas propiedades psicométricas que lo hacen útil para medir la calidad de vida relacionada con la salud en pacientes con insuficiencia cardiaca crónica.

Palabras clave: Insuficiencia cardiaca crónica. Calidad de vida. Validación. Minnesota Living with Heart Failure Questionnaire.
INTRODUCTION

Primary health care (PHC) is the appropriate setting to follow up or patients with chronic heart failure (CHF), as it is a highly prevalent condition that requires frequent monitoring. Few studies have been performed in PHC to investigate the health-related quality of life (HRQOL) of patients with CHF.\(^1,2\) This may hinder the implementation of clinical practice guidelines at that level of care.

The quality of life (QOL) of patients with CHF is more important than is generally recognized.\(^3-6\) In a qualitative study, 49% of patients selected a therapy that would improve QOL, even if it meant a reduction in survival.\(^7\)

Primary care (PC) professionals need a simple and reliable instrument to measure the intervention, treatment, and QOL needs of patients with CHF.

Both generic and specific instruments can be used to measure HRQOL in CHF. One of the most frequently used generic measures is the Short Form-36 Health Survey (SF-36),\(^8,9\) the Spanish version of which has demonstrated good reproducibility and validity.\(^10\) Disease-specific instruments measure aspects of health affected by the condition and are expected to be more sensitive to changes in clinical condition. The most frequently used questionnaires are the Chronic Heart Failure Questionnaire, the Kansas City Cardiomyopathy Questionnaire, and the Minnesota Living with Heart Failure Questionnaire (MLHFQ).\(^11\) The latter is the most widely used in studies of the QOL of patients with CHF in hospital settings,\(^1,2\) and has been shown to have good reliability and validity.\(^12,13\)

The aim of the present study was to evaluate the applicability, internal consistency, and validity of the MLHFQ when used to measure the QOL of CHF patients in primary care, and compare it with the SF-36 questionnaire.

METHODS

The study used a cross-sectional design. To be included, patients were required to have a documented diagnosis of CHF in their PHC clinical records. Diagnosis could be based on a hospital discharge report from a medical specialist or a medical diagnosis made by a PC physician and should be based on Framingham clinical criteria (2 major criteria or 1 major and 2 minor) as well as additional tests such as electrocardiogram, chest x-ray, and echocardiography, with characteristic signs of CHF. A further inclusion criterion was that patients should have been seen by a PC physician in León during at least the prior 6 months. Patients whose heart failure was attributable to a reversible cause were excluded.

Sampling Design

The sample was designed for a larger study (the LEONIC study) in the Leon Health Area, the main objective of which was to assess the epidemiological, clinical, and HRQL characteristics of patients with CHF.

All doctors in the area’s PHC teams were invited to participate. Each professional reviewed paper and computerized records and produced an anonymous register of all patients on their lists with a documented diagnosis of CHF. Computerized records were located using search terms such as “chronic heart failure” and “congestive heart failure.” From the register produced, a proportional sample based on urban and rural strata was selected using systematic random sampling. The sampling interval was defined by the expression: \(N / n\), where \(N =\) census population with CHF and \(n =\) predetermined sample size.

Sample Size

Sample size was calculated to meet the objectives of the LEONIC study, i.e., to estimate the prevalence of clinical features in patients with CHF in both rural and urban areas, with a 95% confidence interval and 4% overall accuracy. It was estimated that a total sample size of 630 patients would be required.

Fieldwork was conducted from January to April 2009. Each patient completed the SF-36 and MLHFQ before being examined by the clinician. A nurse was present to assist if there were problems completing the questionnaires. Clinical and treatment variables, hospital admissions, and any use of the emergency department were also recorded. Functional capacity was assessed using the New York Heart Association (NYHA) scale.

Instruments to Measure Quality of Life

The SF-36 is a validated, self-administered questionnaire for both the general population and patients with a range of diseases. We used the Spanish version 2 of the questionnaire, which includes 36

ABBREVIATIONS

CHF: chronic heart failure
HRQOL: health-related quality of life
NYHA: New York Heart Association
PHC: primary health care
questions measuring 8 dimensions of health: physical function, role-physical, bodily pain, general health, vitality, social function, role-emotional, and mental health. Scores can be summarized in 2 summary components assessing physical and mental health. Items in each dimension are coded, aggregated, summed, and transformed into a scale ranging from 0 (worse health) to 100 (best health).14

Table 1 shows the number of items in each dimension and provides a brief description of the meaning of the scores.

The instrument asks about the extent to which illness has prevented the respondent from living as they would have liked over the last month. The questions cover signs and symptoms of the disease, social relations, physical activity, and sexual, work and emotions topics.

**Statistical Analysis**

Descriptive indices were calculated for both questionnaires. Acceptability was assessed by calculating the percentage of patients with a missing response on any item in each dimension. Ceiling (the percentage of patients obtaining the highest possible score) and floor (the percentage of patients obtaining the lowest possible score) effects were also calculated for each component. Internal consistency was assessed by determining the Cronbach alpha coefficient, which measures homogeneity among items in a given dimension. A value of 0.7 is usually taken to be acceptable for this coefficient for between, group comparisons.21

The 3 aspects of the MLHFQ’s construct validity studied were structural, convergent, and divergent validity. The Kaiser-Meyer-Oklin test (acceptable at values over 0.5) and the Bartlett sphericity test were used to test the statistical adequacy of the factor analysis. Structure was evaluated using exploratory factor analysis with varimax rotation in which each dimension (physical and emotional) was expected to behave as a distinct factor, with item factor loadings over 0.5 and the Bartlett sphericity test were used to test the statistical adequacy of the factor analysis. Structure was evaluated using exploratory factor analysis with varimax rotation in which each dimension (physical and emotional) was expected to behave as a distinct factor, with item factor loadings of 0.4 or more.22

The convergent and divergent validity of the MLHFQ were assessed using Pearson’s multitrait-multimethod correlation matrices to determine whether dimensions measuring similar aspects in the MLHFQ and SF-36 questionnaires correlated better than dimensions that were not conceptually related. A priori hypotheses were that the highest correlations would be observed between the MLHFQ physical component and the physical function, role-physical, and vitality dimensions of the SF-36 and between the emotional component of the MLHFQ and the role-emotional and mental health dimensions of the SF-36. Strong correlations were defined as ≥0.6 and were assumed to indicate convergent construct validity. As regards divergent validity, we hypothesized that correlations would be low (≤0.4) between the SF-36 physical function, role-physical, and vitality dimensions and the MLHFQ emotional dimension and between the physical dimension of MLHFQ and the SF-36 role-emotional and mental health dimensions.

The capacity of the MLHFQ to discriminate between relevant categories of patients was assessed by determining whether being female, having been admitted to hospital, having gone to the hospital emergency room, being in a higher NYHA functional class, and having symptoms of dyspnea or fatigue in the past month led to higher mean scores on the MLHFQ dimensions and lower scores on the SF-36 summary components. We calculated confidence intervals, and statistical testing was performed using...
the Student t test and ANOVA. All analyses were performed using SPSS v. 14.0.

Ethics

The study was approved by the Ethics Committee of Hospital de León on January 27, 2009. Participants were informed of the study objectives and provided signed consent to participate.

RESULTS

A total of 97 physicians participated in the study (46.6% from rural practices). Physicians included 2047 patients aged over 39 years with CHF in the study register (58% female). Of the 630 patients selected to participate, 589 (93.6%) were studied in the office or at home, 11 had died, and 30 did not attend the study visit or were not located. In total, 544 valid questionnaires were obtained for the MLHFQ and 542 for the SF-36. Questionnaires were not answered by 21 patients with moderate or severe dementia and by 14 who were immobilized. A further 10 questionnaires were excluded because of inconsistent data or because of an excessive number of missing responses.

Table 2 shows the demographic and clinical characteristics and QOL summary scores for the patients studied and for patient sub-groups, based on who made the diagnosis. The diagnosis was made at the hospital or by a specialist in 407 (74.8%) patients, and some significant differences were found with the sub-group of patients diagnosed in primary care. In the latter group, patients were older, with a higher proportion of women and those with an etiology attributed to hypertension. No significant differences were observed between the groups in terms of the SF-36 summary components or the MLHFQ dimensions.

Acceptability

Score distributions and internal consistency coefficients for both questionnaires are presented in Table 3. The percentage of participants with a missing response on any item was small, both for MLHFQ dimensions and on the SF-36. Missing responses were somewhat higher for the MLHFQ total score (12.5% with a missing response on any item). This was largely due to items 8 and 10, which measure limitations in professional or sexual activity, respectively, and which are not assigned to any dimension.

The percentage of patients with the worst and the best possible score (floor and ceiling effects, respectively) was low on the MLHFQ, except in the emotional dimension, which showed a ceiling effect of 15.4%. The SF-36 showed substantial floor effects on the role-physical and mental health dimensions, and 4 of its dimensions showed a ceiling effect (over 20% of respondents with the highest possible score).

Internal consistency

All dimensions of both questionnaires showed high internal consistency (α > 0.7), except for the general health dimension (α = 0.68) of the SF-36 (Table 3).

Construct Validity

A value of 0.92 on the Kaiser-Meyer-Olkin test indicated adequate correlation matrices. The Bartlett sphericity test was significant at χ² = 4.213 (P < 0.001), indicating the presence of significant correlations and reinforcing the relevance of the factor analysis. Exploratory factor analysis with two-factor varimax rotation explained 65.8% of the total variance. The 8 items in factor 1 (items 2, 7, 12, and 13) coincided with the items assigned to the physical dimension in the theoretical model and presented high factor loadings in excess of 0.4. On the other hand, item loadings were low for factor 2 (Table 4). The same finding was observed for items 17-21, which showed high factor loadings on factor 2 and low loadings on factor 1. The results reveal two underlying QOL dimensions. In the first factor, item 1, with a saturation of 0.48, increased the amount of variance explained by 8.2%. This item is not assigned to any of the dimensions in the original theoretical model. Other items not assigned to any dimension did not exceed factor loadings of 0.4.

Table 5 shows the high correlations observed between similar domains in both questionnaires. Correlation coefficients between the physical dimension of the MLHFQ and the physical component summary (PCS), physical function, role-physical, bodily pain and vitality dimensions of the SF-36 (column I) ranged from -0.43 to -0.73 and indicate convergent validity (the correlations are negative because the dimensions are oppositely scaled). Moderate correlation coefficients (ranging from -0.5 to -0.55, see column V) were observed between the emotional component of the MLHFQ and the role-emotional and mental health dimensions of the SF-36. In contrast, correlations between dimensions reflecting different domains, which were calculated to assess divergent validity (columns II and IV), were not as low as expected. Values were between -0.33 and -0.51, indicating stronger evidence for convergent validity.

Discriminant Validity

Discriminant validity was studied by determining whether the dimensions of the MLHFQ and the
SF-36 summary component scores were able to distinguish between patients with different levels of CHF severity (Table 6). Women had lower scores than men on both questionnaires ($P<.001$). There was also a progressive deterioration of QOL with increasing NYHA functional class, with mean scores on the physical dimension of MLHFQ ranging from 6.3 for patients in class I to 30.2 in class IV ($P<.0001$). Patients with frequently occurring dyspnea and/or fatigue had statistically significant ($P<.0001$)
The results indicate that the MLHFQ has adequate metric properties in terms of acceptability, internal consistency, and convergent validity. The results of exploratory factor analysis, as in the original model, extracted two factors that explained a significant proportion of the variance. The results indicate the value of the instrument as a measure of HRQOL in patients with CHF who are usually monitored by a primary care physician.

Most of the studies which have measured the QOL of CHF patients in Spain with the MLHFQ or a generic questionnaire have been performed in patients from a hospital setting. The instrument showed acceptable sensitivity to change and good reliability and validity, but these results may not be generalizable to CHF patients who are usually seen in PHC.

The proportion of patients who completed the questionnaire was achieved with minimal support from nurses and was similar to other studies. The low ceiling and floor effects on the MLHFQ questionnaire, compared with the SF-36, and the use of the full range of scores may suggest that the content of this questionnaire better reflects the poorer scores compared to those who were usually symptom-free. The mean scores of those admitted to hospital or attending an emergency department in the past year also reflected poorer QOL than those who were not admitted or had not gone to the emergency department.
specific QOL concerns of these patients, and may point to a greater ability to detect improvement or deterioration.

Internal consistency was satisfactory in all dimensions, with Cronbach alpha coefficients close to or above 0.7, as recommended for group comparisons; these results provide further support for the factor structure emerging from the scale.\textsuperscript{21,30} The coefficients are similar to those of the original version\textsuperscript{17} and to those obtained by Garin et al.\textsuperscript{13} They are also comparable to or higher than the equivalent coefficients for the SF-36, and confirm the reliability of the Spanish version of the MLHFQ.

Evidence for convergent validity was provided by the strong correlations between dimensions of MLHFQ and SF-36 measuring similar concepts, a finding which also provides additional support for the underlying constructs. However, divergent

### TABLE 5. Pearson Correlations Between Minnesota Living With Heart Failure Questionnaire and Short Form-36 Dimensions

<table>
<thead>
<tr>
<th>SF-36</th>
<th>MLHFQ Physical Dimension</th>
<th>MLHFQ Emotional Dimension</th>
<th>MLHFQ Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS</td>
<td>-0.73</td>
<td>-0.44</td>
<td></td>
</tr>
<tr>
<td>Physical function</td>
<td>-0.68</td>
<td>-0.5</td>
<td>-0.65</td>
</tr>
<tr>
<td>Role-physical</td>
<td>-0.58</td>
<td>-0.47</td>
<td>-0.59</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>-0.43</td>
<td>-0.33</td>
<td>-0.44</td>
</tr>
<tr>
<td>Vitality</td>
<td>-0.67</td>
<td>-0.51</td>
<td>-0.65</td>
</tr>
<tr>
<td>IV</td>
<td>V</td>
<td>V</td>
<td>VI</td>
</tr>
<tr>
<td>MCS</td>
<td>-0.44</td>
<td>-0.6</td>
<td></td>
</tr>
<tr>
<td>Role-emotional</td>
<td>-0.4</td>
<td>-0.5</td>
<td>-0.45</td>
</tr>
<tr>
<td>Mental health</td>
<td>-0.45</td>
<td>-0.55</td>
<td>-0.52</td>
</tr>
</tbody>
</table>

**Notes:**
- Abbreviations: PCS, physical component summary; MCS, mental component summary; MLHFQ, Minnesota Living with Heart Failure Questionnaire; SF-36, Short Form-36.
- Correlation significant at 0.01 (2-sided).

### TABLE 6. Mean Scores on Minnesota Living With Heart Failure Questionnaire Dimensions and Short Form-36 Summary Scores According to Patient Characteristics

<table>
<thead>
<tr>
<th></th>
<th>MLHFQ</th>
<th>SF-36</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Physical Dimension</td>
<td>Emotional Dimension</td>
</tr>
<tr>
<td></td>
<td>PCS (95% CI)</td>
<td>MCS (95% CI)</td>
</tr>
</tbody>
</table>

**Notes:**
- Abbreviations: PCS, physical component summary; MCS, mental component summary; MLHFQ, Minnesota Living with Heart Failure Questionnaire; NS, non-significant; SF-36, Short Form-36 Health Survey.
- Data are expressed as means (95% confidence interval).
validity was not completely confirmed, as the correlations between dimensions or components measuring different concepts were not as low as expected.

The strategies used to assess construct validity showed that the scores of each dimension measured the appropriate concepts. Factor analysis extracted 2 factors, indicating a similar structure to that of the original, but our data showed a discrepancy, in the sense that factor loadings for item 1 (‘Has your heart failure caused any swelling of the ankles, legs, etc?’) suggest it could form part of the physical dimension. Construct validation is considered to be the first step in evaluating a new instrument or when there is limited experience with an instrument in a given environment. A study on a small sample of patients who had undergone heart valve surgery confirmed that MLHFQ items fitted well to the physical and emotional dimensions, with factor loadings which were higher than those obtained in our study. This may have been due to the greater similarity between patients included in that study.

In relation to other variables that measure different stages of disease severity, our results were consistent with the a priori hypotheses. MLHFQ scores clearly discriminated between patients in different NYHA classes as well as between those with symptoms of dyspnea and/or fatigue and those admitted to hospital or attending an emergency department in the previous year. In comparison to the SF-36, the MLHFQ appeared to discriminate better between those admitted to hospital or who visited an emergency department and those who did not. These results provide favorable data on criterion and predictive validity and mirror findings from studies performed in non-PC patients. Parajón et al. found that increasing NYHA functional class and hospital admissions during the previous year were associated with higher scores on the MLHFQ, i.e., with poorer QOL. Morcillo et al. also reported good correlations between the MLHFQ, functional class, and the SF-36 in a small sample of patients with advanced CHF.

One of the study limitations concerned the inclusion criteria used to define the study population, as CHF diagnosed by clinical criteria alone can lead to false positives. This possible bias could affect external validity, in the sense that our population may not be representative of the overall population of patients with the condition. However, in our sample, 75% of patients were diagnosed in hospitals and, for those diagnosed in PHC, ultrasound values were available for a substantial proportion (29.2%). The average time from diagnosis was also over 3 years, which would imply that the PC physicians had considerable knowledge of the patient’s clinical condition. For those reasons, we consider selection bias to be unlikely and believe that the sample is probably representative of patients diagnosed with CHF who are monitored by PC physicians. This is one of the study’s main strengths, as the psychometric characteristics of the questionnaire being validated are enhanced by the variety of patients included.

Another limitation of the study is that the measurement properties of reproducibility and sensitivity to change were not investigated. This was due to the study design, although it would be interesting to evaluate these properties in different populations and contexts to help extend use of the questionnaire.

As regards the instrument’s interpretability (i.e., the degree to which a clear meaning can be assigned to scores on an HRQOL instrument), the supposed lack of interpretability of QOL scores such as those of the MLHFQ is a consequence of their novelty or of a lack of widespread use. MLHFQ scores are not standardized, although overall mean scores of 27.7 and 42.7 correspond to NYHA functional classes II and III, respectively, in stable patients, and a change of more than 5 points in total score is considered clinically significant.

To improve interpretability, further studies are needed with the MLHFQ in diverse groups with different treatment needs and where it is possible to predict certain life events.

CONCLUSIONS

This study provides further information on the psychometric properties of the Spanish version of the MLHFQ. It also provides evidence of its potential for future use in both research and in clinical practice in CHF patients seen in primary care, for which there is a shortage of relevant data in the literature.

REFERENCES