Reliability and validity of the Chronic Liver Disease Questionnaire (CLDQ) in adults with non-alcoholic steatohepatitis (NASH)

Kashmira S Chawla,1 Jayant A Talwalkar,2 Jill C Keach,2 Michael Malinchoc,3 Keith D Lindor,4 Roberta Jorgensen,2

ABSTRACT

Introduction: Significant impairments in health-related quality of life (HRQL) in patients with non-alcoholic fatty liver disease have been previously described. The disease-specific HRQL among patients with non-alcoholic steatohepatitis (NASH), however, remains unknown. Aim: To determine the degree of construct validity of the Chronic Liver Disease Questionnaire (CLDQ) in adults with NASH. Methods: Participants referred for the evaluation of hepatology-proven NASH at Mayo Clinic, Rochester, between 1996 and 2000, were evaluated. HRQL assessment by the Short-Form 36 (SF-36) Health Survey and CLD) was performed. The primary outcome was to determine the level of correlation between overall and subscale scores for the CLDQ and SF-36 instruments. Results: Among 79 participants (70%) with NASH completing both questionnaires (mean age, 51.2 years with 64% female gender), excellent reliability was noted for the CLDQ instrument. Significant reductions in all SF-36 domains (p<0.05 for all) including PCS and MCS scores (p<0.02 for both) among participants with NASH compared with normative data from an age-matched and sex-matched US general population sample was observed. Highly significant correlations were observed between overall CLDQ score with SF-36 PCS (r=0.82, p<0.0001) and SF-36 MCS (r=0.67, p<0.0001) scores. Similar degrees of correlation were observed between relevant subscales of the CLDQ and SF-36 as well. Discussion: The CLDQ has excellent reliability and validity of construct for HRQL assessment in adults with NASH when compared with the SF-36. Future investigations among participants with NASH require assessing the responsiveness of the CLDQ to medical therapies and disease progression.

INTRODUCTION

Non-alcoholic steatohepatitis (NASH) is considered as part of a spectrum of liver diseases known as non-alcoholic fatty liver disease (NAFLD). NASH is histologically characterised as the combination of hepatic steatosis with features of inflammation and/or fibrosis.1–4 Obesity, diabetes mellitus and dyslipidemia have been acknowledged as predisposing risk factors for NASH.2–4 Despite its widespread existence, there is currently no widely recognised effective medical therapy for NASH to halt disease progression.5 As a result, the presence of NASH has been associated with the development of cirrhosis and end-stage liver disease in up to 25% of identified cases.6 7 Many participants, however, will continue to be identified with

What is already known about this subject?

▸ Nonalcoholic steatohepatitis (NASH) is a chronic liver disease under the umbrella of non-alcoholic fatty liver disease (NAFLD).
▸ Previous studies have demonstrated that disease-specific Chronic Liver Disease Questionnaire (CLDQ) is a valid psychometric tool to evaluate changes in quality of life among patients with NAFLD.
▸ Patients with NASH experience impairments in multiple health-related quality of life (HRQL) domains as compared to patients with fatty liver alone.

What are the new findings?

▸ CLDQ is a reliable and valid instrument to measure HRQL among patients with NASH.
▸ Both CLDQ and SF-36 scores revealed reductions in physical and mental domains of HRQL, these changes were correlated and significant.
▸ While CLDQ scores were independent of age, sex, and BMI, the presence of diabetes mellitus was associated with reduced physical function.

How might this impact on clinical practice in the foreseeable future?

▸ CLDQ can be used in the NASH patient population to determine treatment response and prognosis.
Among individuals with chronic liver disease including NASH, significant impairments in overall health-related quality of life (HRQL) have been recognised. HRQL is a multidimensional construct based on an individual’s perception of their physical, mental and emotional status. Generic health status instruments such as the Short-Form 36 (SF-36) Health Survey have allowed for comparisons between different chronic disease states and normative data from general populations. Disease-specific instruments, however, can allow for the detection of small but important changes in HRQL over time often missed by generic questionnaires. Recent studies have demonstrated further decrements in HRQL with NASH as compared to fatty liver alone, while the disease-specific Chronic Liver Disease Questionnaire (CLDQ) was evaluated in participants with NAFLD and found to have sound psychometric and discriminant properties for measurement use. The aim of this investigation was to determine the level of construct validity for the CLDQ in a cohort of adults with a diagnosis of NASH.

METHODS AND ANALYSIS

Patient population
Participants referred for the evaluation of NASH at Mayo Clinic, Rochester, between 1996 and 2000, comprised the study cohort. The diagnosis of NASH required (1) abnormal serum liver tests for ≥3 months; (2) liver histology revealing >10% steatosis and lobular inflammation with or without fibrosis; (3) the exclusion of alternate aetiologies for chronic liver disease including chronic viral hepatitis, autoimmune liver disease, primary biliary cirrhosis, chronic biliary obstruction, haemochromatosis, Wilson’s disease, and α1-antitrypsin deficiency; (4) a history of alcohol consumption <40 g/day (men) or <30 g/day (women); and (5) no clinical or biochemical evidence for cirrhosis. All participants underwent a history and physical examination and abdominal ultrasonography to exclude the presence of features consistent with cirrhosis, portal hypertension and biliary obstruction.

Data collection
Demographic and clinical information were abstracted from medical records after Institutional Review Board approval. Demographic variables included participant’s age, gender and body mass index (BMI). BMI was calculated as the ratio of weight in kilograms/(height in metres^2). Clinical variables included serum liver biochemistries (aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (AP), total bilirubin (TB), albumin) and prothrombin time; autoantibodies (antinuclear antibody, antismooth antibody, antimitochondrial antibody); hepatitis serologies (hepatitis B surface antigen, hepatitis B surface antibody, hepatitis C antibody, or hepatitis C RNA level); iron studies (transferrin saturation, ferritin); ceruloplasmin; and α1-antitrypsin levels with phenotype. A history of hypertension was defined by a blood pressure ≥160/90 mm Hg at the initial visit and/or current antihypertensive therapy. A history of type 2 diabetes mellitus defined by a fasting blood glucose ≥126 mg/dL, haemoglobin A1C ≥6%, glycosylated haemoglobin ≥7%, and/or current hypoglycaemic therapy. A history of dyslipidemia (defined as a fasting serum triglyceride level ≥150 mg/dL, low-density lipoprotein (LDL) ≥130 mg/dL, and/or high-density lipoprotein ≤30 mg/dL).

Liver histology
A liver biopsy was performed and made available for assessment in all participants. Specimens were prepared with H&E, Mason’s trichrome and rhodamine stains. Iron stains were also performed. All biopsies were interpreted by hepatopathologists using the Brunt scoring system and without knowledge of the patient’s clinical and biochemical data. The degree of hepatic steatosis was graded on a scale of 1–3: 1=mild (10–30% hepatocyte involvement); 2=moderate (30–70% involvement); 3=severe (>70% involvement). The degree of hepatic inflammation was graded on a scale of 1–3 (mild, moderate, severe). The extent of fibrosis was determined using a four-point scale: 0=None; 1=pericellular and/or pericentral fibrosis; 2=septal or bridging fibrosis and 3=cirrhosis.

Health status instruments
SF-36 health survey
The SF-36 health survey questionnaire was developed for use in the Medical Outcomes Study by the RAND Corporation to study variations in physician practice and patient outcomes. This instrument has been used extensively for clinical research and health policy evaluations. Population norms for individuals residing in the USA, and participants affected by chronic disease states are available. The 36 items measure eight dimensions or domains, including physical function (ten items), social functioning (two), role limitations due to emotional problems (four), role limitations due to physical problems (three), mental health (five), energy/vitality (four), pain (two) and general health perception (five). A single item examines perceptions of health changes in the preceding 12 months and is also included. Item responses vary from dichotomous (yes/no) to a six-point Likert scale ranging from ‘none’ to ‘very severe’. The scoring algorithm includes a summation of item scores for each of the eight domains. Each subscale is transformed before summation. The total score is based on a scale from 0 (poor health) to 100 (excellent health). Physical and mental summary scores may also be derived from the eight domains. The summary score range for physical function is between 8 and 73, while for mental function the range is between 10 and 74.

We also compared HRQL among patients with NASH to ratings from the US general population (N=2474). This group, which constitutes the normative US general
population data for the SF-36, was selected from a random sample of adults without known chronic disease. Individuals from the study were primarily women (61%), married (55%), Caucasian (78%) and completed at least 12 years of education (86%). The mean age of this sample population was 46 years.

**Chronic Liver Disease Questionnaire**

The Chronic Liver Disease Questionnaire (CLDQ) was developed as an evaluative instrument to measure longitudinal change in health status within individuals with chronic liver disease. In addition to measuring both physical and mental health, the instrument was designed to be a disease-specific tool for assessing areas of function important to patients with chronic liver disease. The final version of the instrument contained 29 items contained within six domains including abdominal symptoms (items 1, 5, 17), fatigue (items 2, 4, 8, 11, 13), systemic symptoms (items 3, 6, 21, 23, 27), activity (items 3, 6, 21, 23, 27), emotional function (items 10, 12, 15, 16, 19, 20, 24, 26) and worry (items 18, 22, 25, 28, 29). A Likert scale response format was used for all items ranging from 1 (most impairment) to 7 (least impairment). Scoring of the questionnaire was performed by dividing each domain score by the number of items per domain. Overall CLDQ score was obtained by adding scores for each item and dividing by the total number of items (n=29). Data from study subjects in this cohort were also compared with normative data for healthy controls published elsewhere. Overall reliability (or internal consistency) for the SF-36 and CLDQ instruments was considered excellent, as the Cronbach’s α values for all subscales were greater than 0.78 for both instruments.

**Statistical analysis**

Continuous data were expressed as means±SE, or medians, when appropriate. Categorical data were expressed as the number of participants (or proportion) with a specified condition or clinical variable. The detection of significant differences for continuous variables between groups was performed using the parametric Student t test. Comparisons between frequency data for significant differences were performed using the χ² or Fisher’s exact test method, where appropriate. To provide a basis for comparison, SF-36 scores from this investigation were examined against normative data from an age-matched and gender-matched population. The Pearson product-moment correlation coefficient was used to measure the degree of correlation between overall and related domain-specific scores from the CLDQ instrument and SF-36 health survey. Significance was set at the 0.05 level (two-sided). Statistical analyses were performed using SAS V8.2 (SAS Institute Inc, Cary, North Carolina, USA).

**RESULTS**

Seventy-nine participants fulfilling diagnostic criteria for NASH who completed both health status questionnaires comprised the study cohort (table 1). The mean age of the participants was 46 ± 11 years (range 19–73 years) with 65% women. Mean serum AST and ALT levels were 55 (range 21–299) and 90 U/L (range 16–540). Mean serum AP was 188 U/L (range 79–235), TB 0.6 mg/dL (range 0.2–2.5) and albumin 4.4 g/dL (range 3.1–5). Mean body mass index (BMI) was 31.4 kg/m² (range 19–59), and 60% of participants had BMI ≥30 kg/m²; 19% had a history of diabetes mellitus, 37% had hypertension and 9% were current smokers. Twelve per cent of participants had no evidence for any of the above conditions including BMI >30 kg/m². Histological findings among participants included no fibrosis in 58% of participants, pericentral/perisinusoidal fibrosis in 38% of participants, and septal or bridging fibrosis in 22% of participants.

**Reliability of CLDQ and SF-36 in NASH**

Reliability (or internal consistency) for the SF-36 and CLDQ instruments was considered excellent, as expressed by the high Cronbach’s α values for all subscales of both instruments. The Pearson product-moment correlation coefficient was used to measure the degree of correlation between overall and related domain-specific scores from the CLDQ instrument and SF-36 health survey. Significance was set at the 0.05 level (two-sided). Statistical analyses were performed using SAS V8.2 (SAS Institute Inc, Cary, North Carolina, USA).

**Figure 1** Mean Short Form-36 subscale scores of patients with non-alcoholic steatohepatitis and normative data from an age-matched and sex-matched US general population sample.

**Table 1** Reliability of CLDQ and SF-36 subscales among patients with NASH

<table>
<thead>
<tr>
<th></th>
<th>Items</th>
<th>Chronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLDQ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS</td>
<td>3</td>
<td>0.78</td>
</tr>
<tr>
<td>FA</td>
<td>5</td>
<td>0.91</td>
</tr>
<tr>
<td>SS</td>
<td>5</td>
<td>0.79</td>
</tr>
<tr>
<td>AC</td>
<td>3</td>
<td>0.78</td>
</tr>
<tr>
<td>EF</td>
<td>8</td>
<td>0.93</td>
</tr>
<tr>
<td>WO</td>
<td>5</td>
<td>0.91</td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>0.91</td>
</tr>
<tr>
<td>SF-36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PF</td>
<td>10</td>
<td>0.94</td>
</tr>
<tr>
<td>RP</td>
<td>4</td>
<td>0.93</td>
</tr>
<tr>
<td>RE</td>
<td>2</td>
<td>0.84</td>
</tr>
<tr>
<td>VT</td>
<td>5</td>
<td>0.89</td>
</tr>
<tr>
<td>MH</td>
<td>4</td>
<td>0.90</td>
</tr>
<tr>
<td>SF</td>
<td>2</td>
<td>0.91</td>
</tr>
<tr>
<td>BP</td>
<td>3</td>
<td>0.92</td>
</tr>
<tr>
<td>GH</td>
<td>5</td>
<td>0.84</td>
</tr>
</tbody>
</table>

AC, activity; AS, abdominal symptoms; BP, body pain; CLDQ, Chronic Liver Disease Questionnaire; EF, emotional function; FA, fatigue; GH, general health; NASH, non-alcoholic steatohepatitis; MH, mental health; PF, physical functioning; RE, role emotional; RP, role physical; SF, social function; SF-36, Short-Form 36; SS, systemic symptoms; VT, vitality; WO, worry.
Highly significant correlations between overall CLDQ score with SF-36 PCS (r=0.76, p<0.001) and MCS (r=0.56, p<0.001) were observed (table 3). Highly significant associations (p<0.001) between the CLDQ-FA domain and SF-36 MCS score (r=0.82, p<0.0001), as well as the CLDQ-FA and CLDQ-SS domains with the SF-36 PCS score (r=0.78 and r=0.73, p<0.001) were observed. Other associations of note between related CLDQ and SF-36 domains were SS and BP (r=0.81), FA and VT (r=0.81, EF and MH (r=0.76), AS and BP (r=0.73), AC and BP (r=0.71), and EF with SF (r=0.72).

All reported values are statistically significant (p<0.05).

Associations between demographic and clinical variables on HRQL
The influence of selected variables on reported HRQL was examined with the intent of hypothesis generation in these subgroup analyses. Age, sex, BMI ≥30 kg/m², and fibrosis stage were not significantly associated with CLDQ total, SF-36 PCS and SF-36 MCS scores, respectively (all with p<0.05). In patients with type II diabetes mellitus, however, a significant reduction in the SF-36 mental component summary score was observed (4.1 vs 45, p=0.04) and CLDQ total score (4.1 vs 5.1, p=0.01) was observed when compared with individuals without type II diabetes mellitus.

Associations between serum liver biochemical tests and HRQL
In addition, no correlation between AP, AST, ALT, TB and albumin, with CLDQ total or SF-36 physical and mental component summary scores was observed (data not shown). A weak correlation between increasing BMI and decreased physical component summary score was observed (r=-0.28, p<0.05).
DISCUSSION
The use of generic instruments, such as the SF-36, provides a means for comparing HRQL among healthy and diseased groups with reliable and valid estimates. As observed in populations with chronic liver disease, the presence of significant reductions in HRQL is also observed in a well-defined cohort of patients with NASH when compared with normative data from the US general population. Recently, the CLDQ was evaluated in participants with NAFLD and found to have sound psychometric properties for measurement use. In this study, we extended initial observations regarding the CLDQ by also observing reductions in total and subscale scores, with the CLDQ instrument in all the subscales corresponding to those within the SF-36. Furthermore, the correlations between overall CLDQ score and SF-36 PCS and MCS subscores, as well as corresponding subscale domains were highly significant. These findings suggest that the CLDQ has excellent cross-sectional construct validity with the SF-36 instrument.

HRQL measured by the SF-36 instrument showed similar impairment in overall PCS and MCS functions among participants with NASH compared to other studies. The PCS scores of individuals with NASH in this study are comparable with the PCS scores reported for patients with non-cirrhotic HCV, the MCS score among patients with NASH is lower than chronic hepatitis HCV patients. The patients with NASH had a greater decrease in overall quality of life (PCS and MCS) compared to HBV patients. Using a liver disease-specific instrument, patients with NASH reported CLDQ score similar to NAFLD, with worse scores for abdominal, fatigue and worry domains. Patients with NASH had greater decrease in overall quality of life compared to patients with HCV and HBV.

Criticisms about the use of generic HRQL instruments in assessing participants with chronic disease are well described. HRQL measured by the SF-36 instrument showed similar impairment in overall PCS and MCS functions among participants with NASH compared to other studies. The PCS scores of individuals with NASH in this study are comparable with the PCS scores reported for patients with non-cirrhotic HCV, the MCS score among patients with NASH is lower than chronic hepatitis HCV patients. The patients with NASH had a greater decrease in overall quality of life (PCS and MCS) compared to HBV patients. Using a liver disease-specific instrument, patients with NASH reported CLDQ score similar to NAFLD, with worse scores for abdominal, fatigue and worry domains. Patients with NASH had greater decrease in overall quality of life compared to patients with HCV and HBV.

Criticisms about the use of generic HRQL instruments in assessing participants with chronic disease are well described. HRQL measured by the SF-36 instrument showed similar impairment in overall PCS and MCS functions among participants with NASH compared to other studies. The PCS scores of individuals with NASH in this study are comparable with the PCS scores reported for patients with non-cirrhotic HCV, the MCS score among patients with NASH is lower than chronic hepatitis HCV patients. The patients with NASH had a greater decrease in overall quality of life (PCS and MCS) compared to HBV patients. Using a liver disease-specific instrument, patients with NASH reported CLDQ score similar to NAFLD, with worse scores for abdominal, fatigue and worry domains. Patients with NASH had greater decrease in overall quality of life compared to patients with HCV and HBV.

Results of our study found that the CLDQ was able to detect reductions in physical and mental function identified by the SF-36. However, these results were not influenced by age or sex. Furthermore, there was no relationship between SF-36 or CLDQ scores with serum liver biochemical parameters. This may be related to the vast majority of participants in our study having mild to moderate histological involvement with NASH, and because we did not recruit participants with established compensated or decompensated cirrhosis.

The impact of obesity on HRQL has been previously described. Declines in HRQL appear to occur in parallel with rising BMI values. A greater impact on physical rather than mental function has been observed with bodily pain as an important factor in reduced HRQL which notably can be independent of BMI. Given the relatively narrow BMI range in our study population, we were unable to identify a significant relationship between HRQL and BMI.

The presence of long-standing diabetes mellitus on HRQL has also been studied using the SF-36. Of note, patients with type II diabetes mellitus who are not taking insulin have been observed to report higher HRQL scores using the SF-36 health survey when compared with patients on insulin for type II diabetes mellitus. In our study, we were unable to examine this factor, as none of the patients were taking insulin for diabetes mellitus.

The conduct of our study raised several questions. Data on HRQL in NASH evaluated at our institution may not reflect these values seen in a community or population-based setting. The absence of patients with more aggressive histological change on liver biopsy in this study may also limit the generalizability of results. Finally, comparisons between normative data and other disease populations could not be adjusted for other important determinants of HRQL including race, education level, or BMI.

In conclusion, the CLDQ is a reliable and valid measurement tool of disease-specific HRQL in adults with NASH. Our study found reduced HRQL in all domains of the CLDQ and the SF-36 instruments compared to normative data, and that correlations between relevant CLDQ and SF-36 domains were highly significant. While CLDQ scores were independent of age, sex and BMI, the presence of diabetes mellitus was associated with reduced physical function. Future studies require longitudinal assessment of the CLDQ in larger populations with NAFLD to determine the relationship between HRQL and treatment response as well as prognosis.

Contributors JAT, JCK, MM, KDL and RJ contributed to the conception, acquisition, analysis and interpretation of the data, and drafting and revision of the manuscript. KC contributed to the drafting and revision of the manuscript. All provided final approval of this version.

Competing interests KC will be participating in an unpaid internship regarding surgery at Medtronic.

Patient consent Obtained.

Ethics approval Approval for this study was obtained from the Mayo Clinic Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license,
REFERENCES