Measuring fatigue in haemodialysis patients: The factor structure of the Chalder Fatigue Questionnaire (CFQ)

Federica Picariello ¹ Rona Moss-Morris ¹ Iain C Macdougall ² Joseph Chilcot ¹

¹Health Psychology Section, Psychology Department, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK ²Department of Renal Medicine, King's College Hospital, London, UK

Corresponding author: Dr Joseph Chilcot (joseph.chilcot@kcl.ac.uk). Health Psychology Section, Institute of Psychiatry, Psychology and Neuroscience King’s College London, 5th floor Bermondsey Wing, Guy’s Hospital Campus, London Bridge, London, SE1 9RT

Key words: Chalder fatigue questionnaire; Chalder fatigue scale; fatigue; psychometrics; measurement; confirmatory factor analysis; dialysis; haemodialysis; kidney failure

Conflict of interest: None declared
Abstract

Introduction: Fatigue is common in haemodialysis (HD) patients, leading to poorer quality of life and patient outcomes. Given the complex and subjective nature of fatigue, and its overlap with sleep disturbances and depression, its measurement represents a challenge. Our aim was to evaluate the psychometric properties of the 11-item Chalder Fatigue Questionnaire (CFQ) in HD patients, including an assessment of the validity of the factor structure, internal reliability and discriminant validity with respect to functional impairment due to fatigue, psychological distress and comorbidity.

Methods: Data were evaluated for psychometric analysis from a published study investigating clinical and psychosocial correlates of fatigue among 174 HD patients. Confirmatory factor analysis was used to determine the factor structure using Weighted Least-Squares with Mean and Variance adjustment (WLSMV) estimation. Mplus 7.3 was used for the analysis.

Results: Mental and physical fatigue factors correlated highly (r=.70, p<.01). A bi-factor model with one general fatigue factor, which incorporated three smaller group factors (mental, physical and weakness) had good model fit. The CFQ general factor explained over 85% of the common variance, had high internal consistency, and showed a moderate correlation with distress and a small association with comorbidity and functional impairment.

Conclusions: The CFQ can be summed up to a total fatigue severity score, representing a composite factor of physical and mental symptoms. Taking into consideration the good psychometric properties of the CFQ and its brief length, it should be used in future studies interested in measuring fatigue severity in HD patients.
Introduction

Haemodialysis (HD) patients experience an array of symptoms notably fatigue. Fatigue not only contributes to poor quality of life but also adverse clinical outcomes. A number of fatigue instruments are currently available, such as the Multidimensional Fatigue Inventory (MFI-20), Fatigue Severity Scale (FSS), Piper Fatigue Scale (PFS), or the vitality subscale of the SF-36. The latter is the most widely used instrument in the dialysis population as a marker of fatigue. However, some debates exist in relation to the comprehensiveness of the concept of vitality, capturing a reduction in energy levels, but not necessarily the negative aspects of fatigue, such as weakness, lack of motivation, and difficulty with concentration. The MFI-20 has been criticized for its poor internal reliability and comprehension difficulties in HD patients. Additionally, both the MFI-20 and the PFS are quite lengthy and may be burdensome for fatigued patients. Given the complex and subjective nature of fatigue, and its overlap with sleep disturbances and depression, its measurement represents a challenge. Great variability exists in instruments used to measure fatigue in dialysis patients, with little informed agreement on what instrument(s) are optimal.

Recently we demonstrated that the variability in self-reported fatigue was associated with both clinical and psychological factors, particularly psychological distress, negative beliefs about fatigue and unhelpful behaviours (avoidance and all-or-nothing behaviours). We used the Chalder Fatigue Questionnaire (CFQ) as our measure of fatigue severity, a tool originally developed in Chronic Fatigue Syndrome. The CFQ consists of 11 items loading onto two dimensions of fatigue severity, namely mental fatigue and physical fatigue. The CFQ has been found to have good clinical validity and internal consistency within individuals with CFS. However, there has been no formal assessment of the validity and consistency of the CFQ within...
patients with advanced kidney disease. Furthermore, the advertised multidimensionality of fatigue scales remains contentious with evidence suggesting that many are in fact unidimensional\(^2\), including our recent evaluation of the CFQ in Multiple Sclerosis which demonstrated that the measure can be suitably characterized by a total score\(^3\).

The objective of this study was to evaluate the psychometric properties of CFQ within a sample of HD patients. This secondary psychometric analysis uses data from our previous study investigating clinical and psychosocial correlates of fatigue in HD patients. Our aims were to evaluate the psychometric properties of the CFQ, assessing its factor structure (including the testing of bi-factor models), internal reliability, and discriminant validity with respect to psychological distress, comorbidity and fatigue related functional impairment.

**Methods**

**Patient sample**

268 haemodialysis patients from the renal service of King’s College Hospital, UK were approached. 174 patients consented and were assessed. The mean age of the sample was 58.9 (SD = 1.6) years and the majority of the sample was male (63.2%). Hypertensive renal disease was the most common primary renal diagnosis (23.5%). Patients had been on dialysis for an average of 50.6 months, (minimum=3 months; maximum=303 months). Further information on the sample, including the inclusion/exclusion criteria are described in detail elsewhere\(^1\).

**Questionnaires**

**The Chalder Fatigue Questionnaire (CFQ)**\(^2\), also referred to as the Chalder Fatigue Scale, is an 11-item questionnaire measuring the severity of physical and mental fatigue on two
separate subscales. Seven items represent physical fatigue (items 1-7) and 4 represent mental fatigue (items 8-11). Each item is scored 0-3; less than usual (0), no more than usual (1), more than usual (2) and much more than usual (3). The ratings of items are added together to calculate the total score (range=0-33). High scores represent high levels of fatigue. Further details regarding alternative scoring for the CFQ is described in the appendix.

**Hospital Anxiety and Depression Scale (HADS)**\(^{25}\), is a 14-item self-report measure of mood in patients with medical illnesses. Seven items relate to anxiety and seven relate to depression. Given the high correlation between depression and anxiety subscales, a total score is more appropriate\(^{26}\), with higher scores indicating greater levels of distress (score range 0-42).

**Work and Social Adjustment Scale (WSAS)**\(^{27}\) is a valid and reliable self-report scale of functional impairment attributable to an identified problem (in this case fatigue). The scale consists of five items that correspond to impairment in work, home management, social activities, private leisure activities and relationships. Each item is rated on a 9-point scale ranging from 0 (not at all a problem) to 8 (very seriously impaired), with high scores indicating greater impairment.

**Comorbidity:** Comorbidity was evaluated using the Charlson Comorbidity Index, where higher scores indicate greater morbidity\(^{28}\).

**Statistical Methods**

Confirmatory factor analysis (CFA) was used to evaluate the factor structure of the CFQ, using Weighted Least-Squares with Mean and Variance adjustment (WLSMV) estimation. One, two factor and bi-factor models of fatigue were tested. In addition we tested a 3-factor model as suggested by findings from a large community sample\(^ {29}\). In the bi-factor models, all 11 items
were loaded onto a general fatigue factor. In addition, items were also loaded onto two smaller group factors – physical fatigue (items 1-7) and mental fatigue (items 8-11). Correlations between each of these latent factors were fixed to zero, to capture the unique variance of the latent factors. Assessment of goodness-of-fit was based on standard structural equation modeling criteria including, confirmatory fit index (CFI) >.95, root mean squared error of approximation (RMSEA) <.08, and the Tucker-Lewis index (TLI) >.95. Reliability of the total and subscale scores was assessed using the omega index, along with an indicator of the saturation of a multidimensional scale by a general factor, omega-hierarchical, for the bifactor models. Discriminant validity between the fatigue factors with other patient reported outcomes (distress, comorbidity and fatigue-related functional impairment) was evaluated using Pearson’s correlation. Analyses were conducted in Mplus 7.3.

Results

A summary of the CFA models tested is shown in table 1. One, two and three factor models demonstrated relatively poor fit. In the two factor model, physical and mental fatigue latent factors correlated highly (r=.70, p<.01). A bi-factor model demonstrated improved fit although the RMSEA was still slightly above .08. Modification indices suggested that a residual correlation between items 6 (less strength in muscles) and 7 (feeling weak) would lead to improved model fit. Since these items both appear to measure weakness these items were loaded onto a third group factor “weakness” with the estimated item loadings set to be equal. This slightly modified bifactor model had good fit based upon satisfactory CFI, TLI and RMSEA indices (table 1, model 5). In this model, item 4 (problem starting things) was removed from the physical fatigue group factor since it did not load significantly (p=.57).
The general factor explained 86% of the common variance between items. The mental, physical and weakness group factors explained only a small amount of common variance – 9%, 4% and 1% respectively. Omega hierarchical for the general factor was .90, suggesting that the total score across all items included in the scale predominantly reflects a general fatigue factor. This indicates that a total score for the scale is a reliable indicator for general fatigue and the physical (including weakness) and mental group factors are saturated by the general factor (Cronbach α=0.91).

The general fatigue latent factor had moderate correlations with psychological distress (HADs; r=.64, p<.01) and fatigue-related impairment (WSAS; r=.62, p<.01) and a weak association with comorbidity (r=.27, p<.01).

Discussion

The primary aim of this study was to evaluate the psychometric properties of the CFQ, among HD patients. Our findings failed to support a two-factor model underlying the CFQ, as originally proposed from the scale development in CFS. Bi-factor models allowed the separation of variance into components related to a general factor, group factors and unique variance. This is a useful method to test the dimensionality of measures.

A bi-factor model, consisting of a general fatigue factor, and three smaller group factors (physical and mental fatigue and weakness) most appropriately fitted the data. Items 6 and 7 correlated highly and appeared to measure something specific to weakness. The model fit was most satisfactory when these items were loaded onto a third group factor (weakness). The three group-factors explained a small amount of variance, whereas the general factor explained over 85% of the common variance. Therefore, whilst the CFQ includes different dimensions of
fatigue, it appears sufficiently unidimensional to warrant use of a total score as a reliable measure of general fatigue severity. These findings support those of others\textsuperscript{23, 33}, including a similar analysis conducted within MS patients using the CFQ\textsuperscript{24}.

With regards to discriminant validity, fatigue severity (general factor) had moderate correlations with distress and functional impairment suggesting some overlapping variance, and a weak association with co-morbidity. This suggests a total fatigue score can reasonably discriminate against these related factors.

A few limitations should be noted when evaluating our findings. Only English speakers completed the CFQ, therefore the data cannot be generalised to other languages or cultures. Furthermore, the psychometric evaluation was restricted to the analysis of structural validity, internal reliability and discriminant validity. Test-retest reliability and sensitivity to change remains unknown in this patient group and therefore needs to be evaluated in future studies.

In conclusion, the CFQ appears to be a suitable measure of fatigue severity, owing to its brief length and good psychometric properties. The separation of physical and mental fatigue symptoms is discouraged due to the high saturation by a general fatigue factor. Rather we suggest that a total sum score provides an appropriate and internally reliable measure of general fatigue symptoms in the dialysis population.

References:

Table 1: Summary of CFA model

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>No of free parameters</th>
<th>Chi-square (df)</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-factor</td>
<td>44</td>
<td>385.1 (44)**</td>
<td>.91</td>
<td>.89</td>
<td>.21</td>
</tr>
<tr>
<td>2</td>
<td>2-factor</td>
<td>45</td>
<td>180.7 (43)**</td>
<td>.96</td>
<td>.95</td>
<td>.14</td>
</tr>
<tr>
<td>3</td>
<td>3-factor</td>
<td>48</td>
<td>120.12 (40)</td>
<td>.97</td>
<td>.97</td>
<td>.11</td>
</tr>
<tr>
<td>4</td>
<td>Bi-factor model</td>
<td>55</td>
<td>80.7 (33)**</td>
<td>.98</td>
<td>.98</td>
<td>.09</td>
</tr>
<tr>
<td>5</td>
<td>Bi-factor model (modified)</td>
<td>53</td>
<td>80.1 (35)*</td>
<td>.99</td>
<td>.98</td>
<td>.08</td>
</tr>
</tbody>
</table>

Root mean squared error of approximation (RMSEA); confirmatory fit index (CFI); Tucker-Lewis index (TLI)
Appendix: Scoring the Chalder fatigue questionnaire (CFQ).

Each of the 11 items of the CFQ can be scored using one of two methods:

1. **Continuous, or Likert scoring** allocates 0, 1, 2, and 3 to the items, with scores from 0 to 33. For example, each item is scored; better than usual (0), no worse than usual (1), worse than usual (2) and much worse than usual (3).

2. **Bimodal scoring** allocates 0, 0, 1 and 1 to each of items, which would give a score of 0 to someone who experienced no fatigue, to a maximum fatigue severity score of 11. For example, each item is scored; better than usual (0), no worse than usual (0), worse than usual (1) and much worse than usual (1). Simply, the bimodal scoring counts how many of these items are in fact experienced.
Highlights

- Fatigue is common in dialysis patients yet challenges surrounding its measurement remain.
- No psychometric evaluation of the Chalder fatigue questionnaire in renal dialysis patients exists.
- We found that a bi-factor model, with one general fatigue factor, which incorporated three smaller group factors (mental, physical, and weakness) had good model fit accounting for 85% of the total variance explained.
- A total score represents a reliable measure of fatigue severity in dialysis patients.