Does the Heterogeneity of Chronic Fatigue Syndrome Moderate the Response to Cognitive Behaviour Therapy? An Exploratory Study

Matteo Cella\textsuperscript{a,b} Trudie Chalder\textsuperscript{a} Peter White\textsuperscript{c}

\textsuperscript{a}Department of Psychological Medicine, Institute of Psychiatry, King’s College London, \textsuperscript{b}Department of Clinical, Educational and Health Psychology, University College London, and \textsuperscript{c}Center for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine, Queen Mary University, London, UK

\textbf{Key Words}
Chronic fatigue syndrome \cdot Cognitive behaviour therapy

\textbf{Abstract}
Background: Chronic fatigue syndrome (CFS) is a heterogeneous condition. A few studies have shown that some independent factors predict outcomes after cognitive behaviour therapy (CBT). Two recent systematic reviews suggest that heterogeneity may moderate treatment outcomes. However, no study has explored whether subgroups of CFS predict response to treatment.

Methods: We used both latent class analysis (LCA) and latent class regression (LCR) to clarify the relationship between subgroups of CFS patients (n = 236), diagnosed using the Oxford diagnostic criteria, and the response to CBT. We measured symptoms, demographics, mood, and cognitive and behavioural responses to illness to define subgroups.

Results: We found 5 latent classes by LCA, which did not differ in the direction of their response to CBT, with all classes showing improvement. In contrast, an exploratory LCR identified 4 latent classes, 1 of which predicted a poor response to CBT, whereas the other 3 predicted a good outcome, accounting for more than 70% of the patients. The negative outcome class was defined by anxiety, pain and being focused on symptoms, particularly weight fluctuations and physical shakiness.

Conclusions: CBT should be offered to all classes of patients with CFS, when defined by these measures. It may be possible to predict a minority group with a negative outcome, but this exploratory work needs replication.

\textbf{Introduction}

Chronic fatigue syndrome (CFS) is likely to be an umbrella term for several different subgroups \cite{1–5}. Studies exploring the heterogeneity of CFS have used latent class analysis (LCA), a statistical method that seeks to identify groups of individuals that share underlying indicators referred to as latent constructs \cite{6}. LCA models describe optimal class assignment for each subject so that most variation in the data can be explained by class assignment alone \cite{7, 8}.

Sullivan et al. \cite{9} first applied LCA using symptoms and demographic characteristics to a clinic population of CFS patients and found that these patients could be characterised by several latent classes. Subsequently White et al. \cite{1} in a primary care sample of patients with corroborated postviral fatigue described 2 latent classes. Following these initial investigations, Sullivan et al. \cite{6} conducted an LCA study on fatigue in a large population-based sample, using demographic and clinical measures. The variability of chronic fatigue symptoms could be reduced...
to 5 homogeneous classes: class 1 named ‘CFS like’ comprised of women with a long fatigue duration, a major depressive episode and widespread pain; class 2 named ‘residuals’ comprised of preponderantly men with few symptoms, but depression and sleep problems; class 3 named ‘rheumatic’ accounted for individuals with significant muscle and joint pain; class 4 ‘depressive’ consisted of a high proportion of individuals with comorbid major depression, and finally individuals in class 5, named the ‘acute physical’ class, had significant sore throat, muscle pain and tender lymph nodes.

Other LCA studies determine different classes, defined by the measures employed [4].

Two recent meta-analyses [10, 11] concluded that cognitive behaviour therapy (CBT) was an effective treatment for CFS with a moderate effect size; a Cohen’s d of 0.48 in Malouff et al. [10] and a standardised mean difference of 0.39 in Price et al. [11]. CBT does not seem to cause harm, when properly delivered, and can lead to remission in a minority of patients [12–14]. CBT is generally associated with some dropouts (approx. 16%) but there is also a considerable variation in treatment outcome. The variability in treatment outcomes may be due to the heterogeneity of CFS [10].

Several factors have been associated with poor response to CBT, including physical illness attributions, treatment-resistant depression [15], certain illness cognitions [16], a passive activity pattern, and focusing on bodily symptoms [17], and a low 24-hour urinary free cortisol [18–20]. Nevertheless, the extent to which the heterogeneity of CFS itself, rather than individual factors, relates to treatment outcome has not been investigated before. This type of investigation may help characterise subgroups of patients which respond differently to CBT. The current study therefore explored the heterogeneity of CFS and investigated the moderating influence that subgroups might have on outcome of CBT.

Methods

Participants
Participants for this study were 236 patients with CFS treated with CBT in a specialist CFS clinic in the UK. All patients were referred from primary care and were initially assessed by a consultant psychiatrist or senior therapist who confirmed the diagnosis of CFS. Inclusion criteria for the study were: (1) minimum age of 18 at intake, and (2) a diagnosis of CFS according to the Oxford criteria [21]. All patients undertook a range of laboratory investigations to exclude alternative medical causes, as recommended by the NICE guidelines [22].

Therapists and Treatment
Therapists were either clinical psychologists or nurses with specialist CBT training, regularly supervised and with at least 2 years’ experience with CBT for CFS patients [23]. The number of CBT sessions ranged from 10 to 15 and were all delivered on an individual basis. The CBT offered included encouraging the patient to establish a consistent approach to activity and then building up activity gradually, the establishment of a sleep routine and addressing unhelpful cognitions related to fear of engaging in activity, perfectionism and high standards.

Measures

Chalder Fatigue Scale
The Chalder Fatigue Scale [24] is an 11-item self-rated scale measuring symptoms of physical and mental fatigue. Each item has 4 response options ranging from ‘less than usual’ to ‘much more than usual’. The scale has been used in a large body of research including randomised controlled trials [25, 26] and has good psychometric properties [27]. We used the Likert scores (i.e. 0, 1, 2, 3), with a maximum score of 33.

Hospital Anxiety and Depression Scale
The Hospital Anxiety and Depression Scale [28] is a self-report questionnaire comprising 14 four-point Likert-scaled items measuring symptoms of anxiety and depression over the past week. Scores for each item range from 0 to 3 and the higher the score the more severe the disorder (score range 0–21 for both depression and anxiety).

Work and Social Adjustment Scale
The Work and Social Adjustment Scale [29] is a 5-item self-rated scale that measures patients’ level of impairment in work, home management, relationships, social and private life. Items are scored on an 8-point Likert scale ranging from ‘very severely impaired’ to ‘not impaired at all’ (score range 0–40). The scale has recently been validated for CFS [30].

Cognitive Behavioural Response Questionnaire
The Cognitive Behavioural Response Questionnaire [31] is a self-rated questionnaire designed to measure patients’ cognitive and behavioural responses to illness. Items are rated on a 5-point Likert scale ranging from ‘strongly disagree’ to ‘strongly agree’. Four cognitive and 2 behavioural subscales can be derived: catastrophising, damaging beliefs, symptom focusing, embarrassment avoidance, all-or-nothing behaviour and avoidance/resting [31, 32].

Symptom Checklist
A list of 27 common somatic symptoms was devised. Patients are asked to state categorically if they had experienced the symptom since their illness began. The list of symptoms included those normally associated with CFS (e.g. pain, sore throat, headaches).

Treatment Outcome
The primary outcome used to evaluate CBT effectiveness was the change in fatigue levels, as measured by the Chalder Fatigue Scale, comparing before and after treatment [24].
**Analysis**

Data coding and descriptive statistics were performed with SPSS version 15. LCA and latent class regression (LCR) were computed using Latent Gold version 4.0 [33].

LCA is a non-parametric statistical method used to identify mutually exclusive subgroups of individuals based on their responses to a set of categorical and/or continuous measures [34]. Cases belonging to the same latent class have a homogeneous response pattern on a number of variables while cases belonging to different latent classes are dissimilar. The assumption of latent class models is that the association among certain observed variables is due to an unobserved latent variable within a finite number of mutually exclusive cases [35, 36].

A number of indices are traditionally used to identify the solution for the number of latent classes that best fit the data (e.g. Bayesian information criterion, Akaike information criterion, consistent Akaike information criterion) [37–41].

LCR is a form of LCA that, like regression, predicts a dependent variable as a function of predictor variables. It differs from regression in that an LCR model calculates non-parametric coefficients for latent classes resulting from the association with the dependent variable (in our case – change after treatment). Different classes therefore have different levels of association with the dependent variable according to the influence that the latent variables have. The Wald statistic was used to indicate the discriminative propriety of each predictor, higher scores indicating better discriminative properties.

**Results**

**Patient Characteristics**

Of the 236 patients recruited, 168 (71%) were female. Ethnicity was preponderantly white (91.5%) and the average age was 40.1 (SD 11.3) years. Patients’ self-report scale responses to a set of categorical and/or continuous measures were used in the LCR models.

**Latent Class Analysis**

We chose the 5-class solution for interpretation as it showed lower fit indices compared to the other solutions. In the 5-class solution, class 1 described 26% of the patients, class 2 21%, class 3 21%, class 4 19% and class 5 13%.

Across all classes, CBT produced a median fatigue improvement of 5 [interquartile range (IQR) = 0–11]. The median fatigue scale improvement for the 5 classes after CBT was: 5 (IQR = –1 to +9) for class 1; 4 (IQR = –1 to +10) for class 2; 9 (IQR = 1–12) for class 3; 6 (IQR = 0–12) for class 4, and 7 (IQR = 0–11) for class 5 (fig. 1). Due to the uneven number of patients and the large variance in fatigue change across latent classes, change in fatigue was analysed with non-parametric statistics. The Kruskal-Wallis test did not detect a significant difference in fatigue change between the 5 latent classes [$\chi^2(4) = 6.57, p = 0.16$]. A follow-up $\chi^2$ test was also conducted to ascertain whether the distribution of patients with a fatigue reduction of 10 or more points and patients with a fatigue increase after CBT was different across the 5 latent classes retrieved. The result showed no significant difference in the distribution of these two subgroups [$\chi^2(4) = 2.2, p = 0.69$].

**Latent Class Regression**

The fit indices indicated that the 4-class solution was the model with the best fit, using the Bayesian information criterion and Akaike information criterion AIC3 as the indicative statistics.

The Wald statistic also shows that the 4-class solution had good discriminative proprieties in relation to the dependent variable; hence individuals belonging to different classes had clearer differences in fatigue change after CBT compared to the 2- and 3-class solution.

In the 4-class solution, class 1 described 27% of the CFS patients, class 2 26%, class 3 25% and class 4 22%. The median fatigue scale improvement for the 4 classes after CBT was: −1 (IQR = −5 to +3) for class 1; +7 (IQR = −2 to +12) for class 2; +4 (IQR = 0–10) for class 3, and +7 (IQR = 3–10) for class 4 (fig. 2). Unstandardised regression coefficients representing the contribution of each of the classes to explain the dependent variable were: −10 for class 1, 4.41 for class 2, 5.75 for class 3 and 5.33 for class 4. These coefficients show the direction and strength of the prediction that a particular class had on outcome.

Beta parameters for some of the predictors entered in the LCR models are presented in table 2. Patients in

### Table 1. Mean score and standard deviation of initial assessment scales

<table>
<thead>
<tr>
<th></th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue scale</td>
<td>22.4 ± 7.19</td>
</tr>
<tr>
<td>WSAS</td>
<td>24.52 ± 9.1</td>
</tr>
<tr>
<td>HAD-A</td>
<td>11.33 ± 4</td>
</tr>
<tr>
<td>HAD-D</td>
<td>10.2 ± 3.3</td>
</tr>
<tr>
<td>CBSQ – symptom focusing</td>
<td>12.9 ± 4.8</td>
</tr>
<tr>
<td>CBSQ – fear avoidance</td>
<td>13.1 ± 4.1</td>
</tr>
<tr>
<td>CBSQ – catastrophising</td>
<td>12.7 ± 4.3</td>
</tr>
<tr>
<td>CBSQ – damaging beliefs</td>
<td>10.7 ± 3.4</td>
</tr>
<tr>
<td>CBSQ – embarrassment</td>
<td>11.8 ± 5.6</td>
</tr>
<tr>
<td>CBSQ – all-or-nothing</td>
<td>13.4 ± 5.6</td>
</tr>
<tr>
<td>CBSQ – avoidance/resting</td>
<td>9.1 ± 4.6</td>
</tr>
<tr>
<td>Symptom frequency</td>
<td>15.5 ± 6.1</td>
</tr>
</tbody>
</table>

WSAS = Work and Social Adjustment Scale; HAD-A = Hospital Anxiety and Depression Scale – anxiety score; HAD-D = Hospital Anxiety and Depression Scale – depression score; CBSQ = Cognitive Behavioural Response Questionnaire.
class 1, which was associated with poor CBT outcomes, reported more frequent weight fluctuation, physical shakiness and pain and had higher anxiety and symptom focusing scores compared to the other classes.

Discussion

This is the first study to use LCA to investigate whether the heterogeneity of CFS affects the response to a routine treatment for CFS such as CBT. The most important result was that there was no subgroup of patients with a different response to CBT, as defined by the measures employed, when we used an analysis that was independent of outcome. However, when we added outcome to help define heterogeneity, using LCR, one class was associated with a poor CBT outcome. This class of patients reported a higher frequency of weight fluctuation, physical shaking and pain than other classes and were more symptom focused and anxious.

Two latent classes obtained with LCR, class 2 and class 4, responded particularly well to CBT showing a median reduction of 7 points on the fatigue scale. Features associated with these 2 classes may help to identify patients that are more likely to achieve better outcome after CBT. In particular, class 2 was defined by loss of strength and higher disability scores and class 4 was defined by high depression scores. Finally, class 3 had a more modest median reduction of 4 points on the fatigue scale. This class was characterised by avoidance and resting, suggesting that CBT may be less effective in these patients. Considered together these results indicate that more that 70% of CFS patients attending a specialist clinic improve by 4 points or more on the fatigue scale. This is in keeping with our published routine clinical outcomes where patients improved by a mean of 4 points on the same scale after CBT [42].

This study has some limitations. Firstly, all the measures used in the analysis were self-rated. Although the large majority of the studies conducted on CFS patients routinely use self-rated measures, it would have been ad-
vantageous to gather clinician-based measures and objective measures of activity such as actigraphy as well [17, 43]. A second limitation is the use of only one measure of effectiveness; disability is generally considered alongside fatigue in studies investigating efficacy of behavioural intervention for CFS. Thirdly, alternative diagnostic criteria may have provided different findings [22], although the Oxford criteria require fatigue as the principal symptom [21], which provides clear and easily applied generalisability. Finally, patients were treated in secondary care, so the results cannot be generalised to primary care.

Studies investigating the effects of heterogeneity in moderating treatment response to CBT are novel and promising. We are aware of only one other recent study applying a similar method to investigate treatment response in eating disorders [44]. Variable response to treatment is commonly reported in intervention studies; however, randomised controlled trials have as their main objective the investigation of treatment effectiveness leaving the question of whom the treatment is effective for largely unanswered. Whether influenced by patients’ heterogeneous clinical characteristics, study entry criteria, the role of intervention provider, centre-specific characteristics, co-therapies or study design, outcome variability is a critical aspect to consider. Understanding heterogeneity may be the key to understand why a treatment that is helpful for many may be ineffective in some and even harmful to others [45, 46]. This could be particularly true for CFS and provide some answers to the many controversies in relation to diagnostic criteria and intervention effectiveness [47]. The use of LCA and LCR in relation to treatment outcome has the potential to improve treatment effectiveness by identifying potentially modifiable and unique sets of characteristics which may be more clinically valid and have a better response to different treatments. Further, investigation of symptom heterogeneity can generate empirically based alternative categories of patients that can be used in controlled studies to evaluate intervention effectiveness [48].

Acknowledgments

We thank the patients for completing the questionnaires, Julia Brown and Mary Ridge for entering the data and all the therapists involved in collecting the data: Katharine Rimes, Mary Burgess, Suzanne Roche, Barbara Bowman, Sue Wilkins, Antonia Dittner and Henry Prempeh Adu Bobi. Trudie Chalder acknowledges financial support from the NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and the Institute of Psychiatry, King’s College London.

References


Table 2. Beta parameters for 10 out of the 38 predictors used in LCR

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
<th>Ranking</th>
<th>Wald</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight fluctuations</td>
<td>6.17</td>
<td>1.09</td>
<td>–1.36</td>
<td>–3.99</td>
<td>1/38</td>
<td>1,307.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Breathlessness</td>
<td>–4.64</td>
<td>–5.38</td>
<td>1.31</td>
<td>2.99</td>
<td>2/38</td>
<td>1,129.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Loss of strength</td>
<td>–0.81</td>
<td>9.33</td>
<td>–1.79</td>
<td>2.22</td>
<td>3/38</td>
<td>679</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Avoidance/resting</td>
<td>0.71</td>
<td>–2.49</td>
<td>6.92</td>
<td>0.24</td>
<td>5/38</td>
<td>551.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HAD-anxiety</td>
<td>2.36</td>
<td>1.04</td>
<td>1.18</td>
<td>–5.32</td>
<td>8/38</td>
<td>447.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Shakiness</td>
<td>2.05</td>
<td>–3.99</td>
<td>–1.17</td>
<td>–1.16</td>
<td>11/38</td>
<td>428.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Symptom focusing</td>
<td>3.78</td>
<td>2.13</td>
<td>0.13</td>
<td>–3.26</td>
<td>16/38</td>
<td>298.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HAD-depression</td>
<td>0.13</td>
<td>–0.61</td>
<td>1.95</td>
<td>3.12</td>
<td>19/38</td>
<td>202.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>WSAS</td>
<td>–1.31</td>
<td>2.57</td>
<td>–0.65</td>
<td>0.26</td>
<td>20/38</td>
<td>173.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Pain</td>
<td>1.76</td>
<td>–1.83</td>
<td>–1.73</td>
<td>–1.3</td>
<td>25/38</td>
<td>87.7</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Predictors shown were selected according to their across-class discriminative propriety and their descriptive values. Importance ranking was conducted according to the Wald statistics. HAD = Hospital Anxiety and Depression Scale.


