Predictors of treatment response trajectories to cognitive behavioral therapy for chronic fatigue syndrome: A cohort study

Peer-reviewed author version

Van Oudenhove, Lukas; Debyser, Soetkin; Vergaelen, Elfi; Claes , Stephan & VAN DEN HOUTE, Maaike (2024) Predictors of treatment response trajectories to cognitive behavioral therapy for chronic fatigue syndrome: A cohort study. In: Journal of psychosomatic research, 187 (Art N° 111954).

DOI: 10.1016/j.jpsychores.2024.111954
Handle: http://hdl.handle.net/1942/44601

Predictors of treatment response trajectories to cognitive behavioral therapy for chronic fatigue syndrome: a cohort study

Lukas Van Oudenhove[a,b,c], MD, PhD, Soetkin Debyser[a], MSc, Elfi Vergaelen[d,e], MD, PhD, Stephan Claes[d,e], MD, PhD, Maaike Van Den Houte, PhD[a,b,f]

[a] Laboratory for Brain-Gut Axis Studies (LaBGAS), Translational Research in Gastrointestinal Disorders (TARGID), Department of Chronic Diseases and Metabolism (CHROMETA), KU Leuven, Leuven, Belgium

[b] Leuven Brain Institute, KU Leuven, Leuven, Belgium

[c] Cognitive & Affective Neuroscience Lab (CANLab), Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire, USA.

[d] University Psychiatric Center KU Leuven, University Hospital Leuven, Leuven, Belgium

[e] Mind Body Research, Psychiatry Research Group, Department of Neurosciences, KULeuven, Leuven, Belgium

[f]REVAL - Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University, Diepenbeek, Belgium

Running head: Predictors of treatment response in CFS

Corresponding author:
Maaike Van Den Houte
Laboratory for Brain-Gut Axis Studies (LaBGAS), Translational Research in Gastrointestinal Disorders (TARGID), KULeuven
ON2 Herestraat 49 - bus 701
3000 Leuven
Belgium
Tel: +32 16 37 46 95
E-mail: maaike.vandenhoute@kuleuven.be

Abstract

Background. The response to cognitive behavioral therapy (CBT) for chronic fatigue syndrome (CFS) varies greatly between patients, but predictors of treatment success remain to be elucidated. We aimed to identify patient subgroups based on fatigue trajectory during CBT, identify pre-treatment predictors of subgroup membership, and disentangle the direction of predictor – outcome relationships over time.

Methods. 297 individuals with CFS were enrolled in a standardized CBT program consisting of 17 sessions, with session timing variable between participants. Self-reported levels of fatigue, depressive, anxiety, and somatic symptoms, perceived stress, and positive affect were collected pre-treatment, and after 3, 10, and 15 sessions. Latent Class Growth Analysis (LCGA) was used to identify subgroups based on fatigue trajectories and baseline predictors of group membership. Cross-lagged structural equation models were used to disentangle predictor-outcome relationships.

Results. LCGA identified four fatigue trajectory subgroups, which were labelled as "no improvement" (23%), "weak improvement" (45%), "moderate improvement" (23%), and "strong improvement" (9%) classes. Higher pre-treatment levels of depressive, anxiety, and somatic symptoms, stress, and lower levels of positive affect predicted membership of the "no improvement" subgroup. Reductions in anxiety preceded reductions in fatigue, while the depressive symptoms – fatigue relationship was bidirectional.

Conclusions. On a group level, there were statistically significant reductions in fatigue after 15 sessions of CBT, with important individual differences in treatment response. Higher pre-treatment levels of anxious, depressive, and somatic symptoms and perceived stress are predictors of lack of response, with reductions in anxiety and stress preceding improvements in fatigue.

Keywords: chronic fatigue syndrome, cognitive behavioral therapy, predictors, cross-lagged panel models

Introduction

Chronic fatigue syndrome (CFS) is a disorder characterized by chronic (> 6 months) and debilitating fatigue not alleviated by rest, significant reductions in activity levels, heightened physical and cognitive fatigability, post-exertional malaise, unrefreshing sleep, and complementary symptoms such as muscle pain, cognitive problems and allostatic intolerance. Prevalence of CFS in adults in Western countries and Asia ranges from 0.34 – 2.52% depending on the recruitment strategy and case definition used [1]. The exact mechanisms underlying CFS remain poorly understood, and no biomarkers or diagnostic tests for CFS have been identified. Accordingly, efficacy of both pharmacological and non-pharmacological treatments is limited [2,3], and current treatment guidelines for CFS are aimed at illness management through a multidisciplinary approach [4].

Most researchers agree that an interaction of biological and psychological factors underlie symptom generation and perpetuation in CFS. Consequently, appropriately delivered cognitive behavioral therapy (CBT) is moderately effective in alleviating symptom burden and improving wellbeing and quality of life in individuals with CFS on a group level [5]. This is achieved by modifying cognitions and behaviours that contribute to symptom perpetuation, e.g. through pacing and energy management, reducing symptom focus and increasing self-efficacy, ultimately increasing activity levels and thereby decreasing perceived fatigue and functional limitations. However, response to CBT is very variable between patients, ranging from some patients reporting no benefit or even worsening of symptoms after CBT, to others no longer meeting diagnostic criteria for clinically relevant fatigue afterwards [6]. This heterogeneity in response to CBT is likely a reflection of the heterogeneity in both somatic symptomatology and underlying pathophysiological processes within the population of individuals with CFS [7,8] and is of utmost importance in the decision process concerning disease management. Identifying which patients respond best – and worst – to CBT is a crucial step in optimizing individually tailored care for individuals with CFS.

While the majority of studies evaluating the efficacy of CBT for CFS focus on calculating treatment response at the group level and do not take individual differences and possible predictors of treatment response into account, some predictors of treatment response have been identified. In general, higher pre-treatment fatigue levels, high pain levels, older age, higher levels of non-fatigue symptoms, more depressive symptoms and longer symptom duration have consistently been associated with worse outcomes in individuals with CFS following CBT [9–16]. One study investigating subgroups in treatment response found that the group of poor responders was characterized by higher levels of symptom focus, anxiety, and pain [8]. Importantly, the studies that looked at predictors of treatment success were limited to measuring these variables only at baseline and using these baseline values as predictors in statistical analyses. However, some of the variables that are thought to be predictive of treatment success (such as depressive or anxiety symptoms) are also subject to change during the course of CBT [5]. Consequently, analyses such as these make it impossible to deduce whether changes in, for instance, depression levels drive changes in fatigue levels or vice versa.

Given the current lack of understanding in the heterogeneity of response to CBT in individuals with CFS and the limited research on dynamic interrelationships between predictors of response and outcomes, the goal of the current study was threefold. Aim 1 was to analyse symptom response during a course of standardized CBT in a group of individuals with CFS. Aim 2 was to identify subgroups based on fatigue trajectories during a course of standardized CBT, and baseline predictors of subgroup membership using latent class growth analysis. Aim 3 was to study the dynamic interrelationships between changes in psychological symptoms and fatigue during the course of standardized CBT in CFS using cross-lagged structural equation models.

Materials and Methods

Patients

In this retrospective study, records of all patients enrolled in the government-reimbursed rehabilitation program for CFS developed by the Multidisciplinary Diagnostic Centre for CFS (MDC/CFS; Leuven, Belgium) in association with the University Hospitals Leuven (Leuven, Belgium) and University Psychiatric Centre KU Leuven (Leuven, Belgium) between January 2015 and June 2019 were investigated. Patients are included in the treatment program if they fulfil the 1994 CDC criteria for CFS [17] and report fatigue, rather than pain, to be their primary symptom. The diagnosis is given by a multidisciplinary team after thorough medical and psychiatric evaluation to exclude identifiable causes for the symptoms. Only individuals between 18 and 65 years old who completed the treatment program and who filled out the questionnaires at baseline *and* at least at one other occasion during the course of the treatment program were included in the analyses for this study. The primary purpose of administering these questionnaire was to report back to the National Institute for Health and Disability Insurance. The patients included in the study gave written informed consent for the use of their data for the abovementioned report, but not for the retrospective secondary data analysis reported in this paper. The study was approved by the Ethical Committee Research UZ/KULeuven (Ref. S67906).

Treatment program

Enrolled individuals followed a group psychoeducation session (one half day) and 17 individual 50-minute government-reimbursed CBT sessions (15 regular sessions + 2 booster sessions). The treatment program was administered by an independent local licensed psychotherapist that followed the program guidelines as developed by the MDC/CFS. These guidelines are publicly available (only in Dutch) at https://www.riziv.fgov.be/SiteCollectionDocuments/cognitief_gedragstherapeut_cvs_therapieprotocol.pdf. Participants of the treatment program were free to choose the time span in which these 17 sessions took place (in agreement with their psychotherapist), but the two booster sessions were always performed at least three months after the final regular session. The general goals of the individual CBT sessions were as follows: dosing and developing of a healthy, variable circadian rhythm adjusted to the individual's capacity; learning how to cope with general and emotional illness perpetuating factors in

order to lessen the physical burden and improve acceptance; and changing important personal or familial illness perpetuating factors. When the first signs of recovery appeared, the focus shifted to increasing physical, social and mental workload. Throughout the process, individual vulnerabilities such as perfectionism or incapability of setting limits/borders were targeted as well.

Measures

Individuals enrolled in the treatment program filled out a questionnaire battery evaluating physical and mental health at first admission to the Multidisciplinary Diagnostic Center, as well as after three (follow-up 1) and ten (follow-up 2) CBT sessions and after the end of treatment (after session 15, before the booster sessions). The following questionnaires were used in the secondary analyses reported in this paper:

- The *Checklist Individual Strength (CIS-20)* [18] consists of 20 items scored on a 7-point Likert scale and measures several aspects of fatigue. Subscales include subjective fatigue experience, concentration, motivation and physical activity levels, with higher scores reflecting higher fatigue severity for all subscales.
- The *MOS 36-item Short Form Health Survey* [19] (SF-36) consists of 36 items measuring health-related quality of life. A physical and mental composite score can be calculated (theoretical range 0 – 100) with higher scores indicating higher health-related quality of life. The composite scores are calculated in such a way that a score of 50 refers to an average healthy individual according to normative values [20].
- The Patient Health Questionnaire-15 (PHQ-15) [21] was used to measure general somatic symptom severity. Respondents indicate to what extent they were bothered by each of the 15 listed symptoms for the past four weeks on a scale from 0 to 2.
- The Patient Health Questionnaire-9 (PHQ-9) [22] was used to measure severity of depressive symptoms. Respondents indicate the frequency of 9 symptoms of depression over the past two weeks on a scale from 0 to 3.
- The Generalized Anxiety Disorder-7 (GAD-7) [23] was used to measure severity of anxiety symptoms. Respondents indicate the frequency of anxiety symptoms over the past two weeks on a scale from 0 to 3.
- The *Perceived Stress Questionnaire (PSQ*) [24] was used to measure daily perceived stress. The PSQ consists of 30 items that are rated on a 4-point Likert scale (range 1 - 4). A perceived stress index can be calculated by linearly transforming the total score to a score between 0 and 1.
- The trait version of the *Positive and Negative Affect Schedule (PANAS*) [25] measures positive and negative affectivity. Respondents indicate how often (range 1 - 5) they experience 10 positive and 10 negative emotions in daily life. Only the positive affect subscale was used in this study.

Statistical analysis

Aim 1. Marginal linear mixed models were used to evaluate the general progress of the patients throughout the course of treatment with regards to the questionnaires listed above. In case of a significant main effect of time, all pairwise comparisons with Tukey-Kramer corrections were performed to specify at which timepoints the scores were significantly different from one another. The main outcome variable of interest was the total score of the CIS-20. For this variable we also investigated the number of patients that showed clinically relevant change in fatigue scores. For this purpose, we calculated the reliable change index for the total score of the CIS-20, which indicates how great reduction in fatigue should be so that it's unlikely to be due to measurement error based on the reliability of the CIS-20[26].

Aim 2. Latent class growth analysis (LCGA) was used to identify different subgroups based on fatigue trajectories throughout the course of treatment using the SAS macro TRAJ [27]. In this data-driven analysis technique, individuals were clustered together based on their CIS-20 score before treatment (intercept) and change in fatigue (total CIS-20 scores) over time (linear and higher-order slopes), allowing the identification of different classes based on fatigue trajectory. The choice of the optimal number of classes was based on 3 criteria: 1) The Bayesian Information Criterion (BIC) of the model with k+1 classes must be lower than the BIC of the model with k classes. 2) Classes needed to contain minimally 5% of subjects. 3) Differences between classes needed to be interpretable. Once the optimal class solution was identified, risk factor analysis was performed to identify whether pre-treatment scores on the PHQ-9, PHQ-15, GAD-7, PANAS and PSQ predicted class membership. Further, we used marginal linear mixed models to investigate whether fatigue trajectory class allocation was related to the changes in physical health-related quality of life (physical composite score of the SF-36) during the course of treatment (results in supplement).

Aim 3. Cross-lagged structural equation models were used to investigate the directionality of effects between changes in the CIS-20 score on the one hand and changes in the predictors of treatment outcome identified in aim 2 (PHQ-9, PHQ-15, GAD-7, PANAS and PSQ scores) on the other hand. Specifically, these models test whether values of variable X at time T predict changes in values of variable Y at time T+1 and/or vice versa (cross-lagged relationships), while controlling for all auto-regressive coefficients (i.e. stabilities over time) and within-time (i.e. cross-sectional) correlations [28]. Five cross-lagged models were investigated with variable X = CIS-20 scores and Y = PHQ-9, PHQ-15, GAD-7, PANAS, and PSQ scores respectively. All models were tested with different combinations of fixed and non-fixed cross-lagged paths, while auto-regressive paths were not fixed. Sensible auto-regressive paths based on modification indices were included to achieve optimal fit. Detailed information on fit indices of the final cross-lagged models can be found in supplementary table 3. The SAS-code of the final models reported in this paper is available at https://github.com/labgas/proj_CFS_treatment and can be used to identify the exact model structure and specifications for all variables.

All analyses were performed with SAS 9.4 (SAS Institute, Cary, NC, USA). Only individuals who had completed the questionnaires before treatment (baseline) and at least one other occasion were included. Missing values were handled in all analyses by (full information) maximum likelihood estimation.

Results

Sample description and missing data

The average age of the respondents was 40.9 years (SD = 9.7). The majority of respondents (n = 172, 57.9%) were not working due to their fatigue symptoms, 118 respondents (39.7%) worked fulltime or parttime, 2 (0.01%) did voluntary work, 3 (0.01%) were students, and 2 (0.01%) were retired. At the start of treatment, 118 respondents (39.7%) were taking antidepressants, 55 respondents (18.5%) were taking benzodiazepines and 35 respondents (11.8%) were taking opioids. The average symptom duration at the start of treatment was 5.2 years (SD = 5.5 years, range 0 – 50 years). The length of treatment (excluding booster sessions) was one average 8.3 months (SD = 2.2 months; range 4.3 – 21.3 months, with 90% of respondents having completed the treatment between 5.9 and 11.7 months).

For an overview of the frequency of completion rates (4, 3, or 2 out of 4 measurement moments) per fatigue trajectory class, see Table 1. For an overview of the completion rates of the questionnaires per measurement occasion, see supplementary table 1. Because the SF-36 and PSQ were not part of the questionnaire battery at follow-up 1 and 2 until the end of 2016, the response rates for these questionnaires are lower. To investigate whether individuals with missing data differed in treatment response (CIS-20 scores) from individuals with complete data, binary variables were created for each questionnaire indicating whether the participant had data for all time points or not. Consequently we performed 7 mixed model analyses (one for each questionnaire) with CIS-20 scores as a dependent variable and time, the binary variable, and the time*binary variable interaction as independent variables. No significant effects of the binary variable were found (all p's > 0.10) indicating that the missingness was not related to our main outcome of interest.

Aim 1: Overall symptom development during course of treatment (Table 2)

As shown in Table 2, there were significant reductions in all aspects of fatigue, general somatic symptoms, depressive and anxiety symptoms, and perceived stress, as well as significant increases in physical health-related quality of life and positive affect at the group level. No changes in mental health-related quality of life were found. Effect sizes were small to medium.

Zooming in on our main outcome variable of interest, the total score of the CIS-20, the average score before treatment was 112.5 (SD = 13.6), which is close to the mean score of the population of individuals with CFS [18]. Although the reduction to 98.1 (SD = 25.5) after treatment was significant, this post-treatment average score is still well above the cut-off for clinically significant fatigue, which is 76 [29] – as shown in Figure 1. After treatment, 81.1% of the sample still experienced clinically

significant fatigue. The reduction in fatigue scores from pre- to post-treatment was considered reliable for 38.7% of individuals according to the Reliable Change Index [26].

Aim 2: Subgroups in fatigue trajectory and risk factors

Descriptives of all tested LCGA models on the total score of the CIS-20 can be found in Supplementary Table 2. The intercept and slope of the four classes in the selected LCGA solution can be found in Table 3 and a visual depiction of the four estimated fatigue trajectories can be found in Figure 2.

Firstly, a non-improvement class, consisting of 23% of patients, could be identified. This class consisted of individuals that started the treatment with very high fatigue levels that did not improve throughout the course of treatments. Individuals in the three other classes had lower (compared to the non-improvement class) and similar (compared to each other) initial fatigue levels, but could be distinguished by the extent to which fatigue reduced over time (i.e. slopes). Amongst these classes an improvement class, that had large reductions in fatigue during the course of treatment, could be identified. However, this class only contained 9% of the included individuals. The remaining two classes consisted of a "moderate improvement" class (23% of included individuals) and a "weak improvement" class (45% of included individuals). Demographics and average questionnaire scores before treatment for the four classes can be found in Table 4. There was a significant difference in age between the classes ($F_{3,296} = 6.12$, $p < 0.001$), with individuals in the non-improvement class being significantly older ($M = 44.77$, $SD = 8.46$) compared to weak ($M = 40.68$, $SD = 9.28$) and moderate ($M = 37.87$, $SD = 11.13$) improvement classes. There were no differences in the distribution of sex between the different classes ($X^2 = 0.54$, $p = 0.91$).

Within the LCGA analysis, we investigated whether depressive, anxiety, and somatic symptom severity, positive affect, and perceived stress before treatment were predictors of class allocation. Because of the significant difference in age between classes, we controlled for age in these analyses. The non-improvement class was used as a reference class. The results are summarized in Table 5. In summary, higher levels of depressive and anxiety symptoms, somatic symptom severity, and perceived stress before treatment increased the chance of being allocated to the non-improvement class, and thus of having little reduction in fatigue throughout the course of treatment. Conversely, higher pre-treatment scores of positive affect significantly decreased the chance of being allocated to the non-improvement class.

Aim 3: Directionality of effects: cross-lagged analyses

An overview of cross lagged models can be found in Figure 3. Cross-lagged analysis showed a significant bidirectional positive relationship between depressive symptoms and fatigue over time, indicating that lower levels of depressive symptoms predicted larger reductions in fatigue over time, while lower levels of fatigue predicted larger reductions in depressive symptoms over time (Fig. 3A). A similar bidirectional, but negative, relationship was found between positive affect and fatigue, indicating

that higher positive affect predicted larger decreases in fatigue over time and lower fatigue predicted larger increases in positive affect over time (Fig. 3B). Lower perceived stress (Fig. 3C) and lower anxiety (Fig. 3D) predicted larger reductions in fatigue over time, though fatigue did not predict changes in either anxiety or stress over time. In contrast, lower levels of fatigue predicted larger reductions in somatic symptom severity over time, but somatic symptom severity did not predict changes in fatigue over time (Fig. 3E).

Discussion

The aims of this study were 1) to analyse symptom response on the group level in individuals with CFS following a standardized CBT course, 2) to identify subgroups based on fatigue trajectories during this CBT course as well as baseline risk factors of poor response, and 3) to disentangle the directionality of relationships between psychological symptoms and fatigue levels during CBT. Questionnaire scores of 297 individuals with CFS, enrolled in a standardized CBT program, were investigated. On a group level we found statistically significant reductions in different aspects of fatigue, somatic symptom severity, anxiety, depressive symptoms, perceived stress and negative affect as well as statistically significant increases in physical health-related quality of life and positive affect during the course of treatment. The effect sizes were small to medium, with a medium-sized effect for the main variable of interest, the total score of the CIS-20, which measures different aspects of fatigue. This is comparable to the effect sizes reported in a recent meta-analysis studying the efficacy of CBT on fatigue levels [5], although effect sizes from this meta-analysis using randomized controlled trials cannot directly be compared to our naturalistic study without control group.

As could be expected based on earlier treatment studies in CFS [6], more than 80% of patients in our study still met the cut-off for clinically significant fatigue after CBT. Figure 1 clearly indicates that inter-individual differences in fatigue levels in this sample became larger throughout the course of treatment, corroborating the heterogeneous response to treatment in this population and highlighting the need to identify predictors of treatment response. To achieve this, respondents were divided in different classes reflecting their fatigue trajectory during the course of CBT. We found that only a small minority of individuals with CFS showed strong reductions in fatigue throughout the treatment course, while almost a quarter of the sample - who also had higher fatigue scores before treatment - did not benefit from the treatment at all, at least in terms of fatigue levels. This finding is similar to findings reported by Cella et al., 2011 [8], who used comparable analysis techniques to investigate subgroups in treatment response. Moreover, our results indicated that higher levels of depressive, anxiety, and somatic symptoms, higher levels of perceived stress, and lower levels of positive affect before treatment are risk factors of poor treatment outcome. These results confirm previous findings on predictors of response to CBT in CFS, showing that patients with higher levels of psychiatric symptoms at baseline have a lower chance at recovery and symptom improvement[10,15,16].

Since depressive symptoms, stress, anxiety and positive affect are expected to change throughout the course of CBT, we carried out cross-lagged panel analysis to disentangle the directionality of effects between reductions in fatigue and the abovementioned "predictors". Interestingly, the cross-lagged analyses revealed that while reductions in anxiety and perceived stress preceded reductions in fatigue, reductions in fatigue preceded reductions in other somatic symptoms. The relationship between depressive symptoms and fatigue was bidirectional, indicating that reductions in fatigue and reductions in depressive symptoms reinforced each other, creating a positive spiral. The same was true for positive affect: increases in positive affect encouraged reductions in fatigue and vice versa.

The components of the CBT protocol that was investigated here are in line with current guidelines for CFS management [4] and similar to CBT programs investigated in earlier prediction studies [8,12]. However, while our results show that 60-70% of the included individuals experience reductions in fatigue (weak, moderate, and strong improvement classes), for a majority of the patients fatigue levels do not end up below the cut-off of clinically problematic fatigue. Figure 1 and figure 2 clearly show that fatigue levels decline at different rates for different (classes) of individuals with CFS, and that there are still meaningful reductions in fatigue in between measurement moments 3 and 4 for a number of individuals. This might indicate that some individuals, particularly those in the weak and moderate improvement classes, may possibly benefit from a longer treatment or more CBT sessions than are currently offered in the investigated program – an observation that is supported by studies showing that CBT protocols that offer more sessions or a higher therapy dosage generally produce better results in CFS [5]. Additionally, other pharmacological and non-pharmacological interventions might be beneficial for specific subgroups of individuals with CFS. While the National Institute for Health and Care Excellence (NICE)[4] guidelines state that there is no definitive cure for CFS, recommended interventions include personalized, supervised physical activity and exercise programs, advice on rest and sleep cycles, pharmacological symptom management, and dietary adjustments.

A minority of the sample (23%) did not benefit from the treatment at all, both in terms of fatigue level (Figure 2) and in terms of health-related quality of life. The fact that individuals with higher depressive symptoms, anxiety, and stress at baseline had a higher chance of belonging to this group highlights the need to assess these factors before the start of treatment and take them into account during treatment individualization. Individuals with high levels of anxiety, depressive symptoms and stress might greatly benefit from incorporating strategies to reduce these factors into the treatment. While this is already the case in the Belgian program to a limited extent – e.g. anxiety is targeted by reducing avoidance strategies, depressive symptoms and stress are targeted by working on emotion regulation strategies and maladaptive cognitions) – there is certainly room for increased focus on individual vulnerabilities and perpetuating factors. If levels of psychological symptoms are so high they hamper progress during the treatment program, these symptoms may need to be targeted even before the individual enters the CFS program, which is primarily aimed at fatigue reduction. For instance,

individuals with high level of depressive symptoms might lack the motivation to apply learned strategies and skills to daily life (although it should be noted that individuals suffering from major depressive disorder are excluded from the program as this is considered a psychiatric cause for the fatigue). Similarly, individuals suffering from (severe) anxiety, whether it be in the form of more generalized anxiety or somatic symptom-specific anxiety, should be treated for this in a specialized environment. Whether psychological vulnerabilities have to be targeted before the start of a program aimed at fatigue reduction or incorporated in it, should be reviewed and decided on in a case-by-case manner, depending on the severity and exact content of the symptom. Additionally, the cross-lagged analysis showed that reductions in stress and reductions in anxiety preceded reductions in fatigue but not vice versa, indicating that it might be beneficial for individuals with high anxiety and stress levels to work on this before starting the CBT program focused on decreasing fatigue levels. Conversely, reductions in fatigue on the one hand and reductions in depressive symptoms and increases in positive affect on the other hand seem to reinforce one another, indicating that it might be beneficial to incorporate them in the program. However, given the observational design of the study and the absence of a no-treatment control group, these findings and their implications should be interpreted with caution.

The current study has several notable strengths, including acknowledgment of the heterogeneity in treatment response, the large sample size, and the use of cross-lagged analysis to disentangle directionality of predictor – outcome relationships. Moreover, while most research tends to focus on risk factors for unfavourable outcomes, our study highlights the potential for positive affectivity as a protective factor. Importantly, this is – to our knowledge - the first study to explore the dynamic interrelationship between psychological predictors and fatigue over time in individuals with CFS undergoing CBT. As outlined above, these analyses provide us unique insights in the importance of addressing the right symptom at the right time. Additionally, the study's robust sample of 297 patients were all diagnosed by a multidisciplinary team after thorough medical and psychiatric evaluation based on consistent criteria, ensuring the integrity of the sample.

The current study suffers from some limitations that should be acknowledged. First of all, this is a secondary analysis on a dataset containing questionnaires filled out by a group of individuals with CFS in the context of a government-funded treatment program. This naturalistic setting has a few consequences: 1) we had no control over the used questionnaires and there were possibly other psychosocial factors that might play a large role in the perpetuation of symptoms (such as trauma, social support, etc) that we could not take into account, 2) there were no standardized intervals between therapy sessions or measurement moments, 3) there was no no-treatment control group. Additionally, although we have information on medication use, this was not controlled for during our analyses due to insufficient statistical power. Further, while this study investigated individual differences in treatment response, analyses were conducted at the group level. Alternative designs such as N = 1 designs might offer more in-depth insights in symptom fluctuation patterns and support the development of tailored interventions. Finally, our measurement of fatigue and its predictors was limited to a unimodal, self-

reported approach. Future research should consider 1) multimodal assessment of fatigue, incorporating ambulatory measures to assess fatigue and activity in daily life, as well as more objective endurance tests, and 2) including a wider spectrum of biopsychosocial predictors to capture the heterogeneity of CFS pathophysiology.

In summary, the current study highlights large interindividual differences in the effect of CBT on fatigue levels in individuals with CFS. The findings suggest that higher psychological symptom levels predict poorer outcomes after CBT, with reductions in anxiety and stress appearing to precede improvements in fatigue. Patients presenting with elevated psychological symptoms may benefit from integrating strategies aimed at addressing these factors as part of their treatment.

Acknowledgements

The authors would like to thank Luscinda Aerts for her help with data collection.

Author contributions

Lukas Van Oudenhove: conceptualization, formal analysis, methodology, supervision, writing – review and editing. **Soetkin Debyser**: data curation, investigation, writing – review and editing; **Elfi Vergaelen**: conceptualization, methodology, writing – review and editing; **Stephan Claes**: conceptualization, methodology, writing – review and editing; **Maaike Van Den Houte**: conceptualization, data curation, formal analysis, investigation, methodology, project administration, supervision, visualization, writing – original draft.

Financial support

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Ethical standards

This study protocol was reviewed and approved by the Ethical Committee Research UZ/KULeuven, approval number S67906. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

[1] E.J. Lim, Y.C. Ahn, E.S. Jang, S.W. Lee, S.H. Lee, C.G. Son, Systematic review and meta-analysis of the prevalence of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME), J Transl Med 18 (2020). https://doi.org/10.1186/s12967-020-02269-0.

[2] B.D. Castell, N. Kazantzis, R.E. Moss-Morris, Cognitive behavioral therapy and graded exercise for chronic fatigue syndrome: A meta-analysis, Clinical Psychology: Science and Practice 18 (2011) 311–324. https://doi.org/10.1111/j.1468-2850.2011.01262.x.

[3] A. Collatz, S.C. Johnston, D.R. Staines, S.M. Marshall-Gradisnik, A Systematic Review of Drug Therapies for Chronic Fatigue Syndrome/Myalgic Encephalomyelitis, Clin Ther 38 (2016) 1263-1271.e9. https://doi.org/10.1016/j.clinthera.2016.04.038.

[4] National Institute for Health and Care Excellence (Great Britain), Myalgic encephalomyelitis (or encephalopathy)/chronic fatigue syndrome : diagnosis and management., n.d.

[5] F. Maas genannt Bermpohl, A.C. Kucharczyk-Bodenburg, A. Martin, Efficacy and Acceptance of Cognitive Behavioral Therapy in Adults with Chronic Fatigue Syndrome: A Meta-analysis, Int J Behav Med (2024). https://doi.org/10.1007/s12529-023-10254-2.

[6] J.M. Malouff, E.B. Thorsteinsson, S.E. Rooke, N. Bhullar, N.S. Schutte, Efficacy of cognitive behavioral therapy for chronic fatigue syndrome: A meta-analysis, Clin Psychol Rev 28 (2008) 736–745. https://doi.org/10.1016/j.cpr.2007.10.004.

[7] E. Aslakson, U. Vollmer-Conna, W.C. Reeves, P.D. White, Replication of an empirical approach to delineate the heterogeneity of chronic unexplained fatigue, Popul Health Metr 7 (2009) 17. https://doi.org/10.1186/1478-7954-7-17.

[8] M. Cella, T. Chalder, P.D. White, Does the heterogeneity of chronic fatigue syndrome moderate the response to cognitive behaviour therapy? An exploratory study, Psychother Psychosom 80 (2011) 353–358. https://doi.org/10.1159/000327582.

[9] E. Crawley, S.M. Collin, P.D. White, K. Rimes, J.A. Sterne, M.T. May, Treatment outcome in adults with chronic fatigue syndrome: a prospective study in England based on the CFS/ME National Outcomes Database., QJM 106 (2013) 555–565. https://doi.org/10.1093/qjmed/hct061.

[10] L. Darbishire, P. Seed, L. Ridsdale, Predictors of outcome following treatment for chronic fatigue, British Journal of Psychiatry 186 (2005) 350–351. https://doi.org/10.1192/bjp.186.4.350.

[11] S. Kempke, L. Goossens, P. Luyten, P. Bekaert, B. Van Houdenhove, P. Van Wambeke, Predictors of outcome in a multi-component treatment program for chronic fatigue syndrome, J Affect Disord 126 (2010) 174–179. https://doi.org/10.1016/j.jad.2010.01.073.

[12] K.M.G. Schreurs, M.M. Veehof, L. Passade, M.M.R. Vollenbroek-Hutten, Cognitive behavioural treatment for chronic fatigue syndrome in a rehabilitation setting: Effectiveness and predictors of outcome, Behaviour Research and Therapy 49 (2011) 908–913. https://doi.org/10.1016/j.brat.2011.09.004.

[13] A. Janse, G. Bleijenberg, H. Knoop, Prediction of long-term outcome after cognitive behavioral therapy for chronic fatigue syndrome, J Psychosom Res 121 (2019) 93–99. https://doi.org/10.1016/j.jpsychores.2019.03.017.

[14] C.X. Sandler, B.A. Hamilton, S.L. Horsfield, B.K. Bennett, U. Vollmer-Conna, C. Tzarimas, A.R. Lloyd, Outcomes and predictors of response from an optimised, multidisciplinary intervention for chronic fatigue states, Intern Med J 46 (2016) 1421–1429. https://doi.org/10.1111/imj.13251.

[15] E. Flo, T. Chalder, Prevalence and predictors of recovery from chronic fatigue syndrome in a routine clinical practice, Behaviour Research and Therapy 63 (2014) 1–8. https://doi.org/10.1016/j.brat.2014.08.013.

[16] L. Sarter, J. Heider, L. Kirchner, S. Schenkel, M. Witthöft, W. Rief, M. Kleinstäuber, Cognitive and emotional variables predicting treatment outcome of cognitive behavior therapies for patients with medically unexplained symptoms: A meta-analysis, J Psychosom Res 146 (2021). https://doi.org/10.1016/j.jpsychores.2021.110486.

[17] K. Fukuda, S.E. Straus, I. Hickie, M.C. Sharpe, M. Psych, J.G. Dobbins, A. Komaroff, The Chronic Fatigue Syndrome: A Comprehensive Approach to Its Definition and Study, 1994. https://annals.org.

[18] M. Worm-Smeitink, M. Gielissen, L. Bloot, H.W.M. van Laarhoven, B.G.M. van Engelen, P. van Riel, G. Bleijenberg, S. Nikolaus, H. Knoop, The assessment of fatigue: Psychometric qualities and norms for the Checklist individual strength, J Psychosom Res 98 (2017) 40–46. https://doi.org/10.1016/j.jpsychores.2017.05.007.

[19] C. McHorney, J. Ware Jr, A. Raczek, The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs., Med Care 31 (1993) 247–263.

[20] J.E. Ware, J. Ware, SF-36 Physical and Mental Health Summary Scales: a User's Manual, 1993. https://www.researchgate.net/publication/292390260.

[21] K. Kroenke, R.L. Spitzer, J.B.W. Williams, The PHQ-15: Validity of a New Measure for Evaluating the Severity of Somatic Symptoms, 2002. http://journals.lww.com/psychosomaticmedicine.

[22] K. Kroenke, R.L. Spitzer, J.B.W. Williams, The PHQ-9: Validity of a brief depression severity measure, J Gen Intern Med 16 (2001) 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

[23] R.L. Spitzer, K. Kroenke, J.B.W. Williams, B. Löwe, A Brief Measure for Assessing Generalized Anxiety Disorder The GAD-7, Arch Intern Med 166 (2006) 1092–1097. https://jamanetwork.com/.

[24] H. Fliege, M. Rose, P. Arck, O.B. Walter, R.D. Kocalevent, C. Weber, B.F. Klapp, The Perceived Stress Questionnaire (PSQ) reconsidered: Validation and reference values from different clinical and healthy adult samples, Psychosom Med 67 (2005) 78–88. https://doi.org/10.1097/01.psy.0000151491.80178.78.

[25] D. Watson, L.A. Clark, A. Tellegen, Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales, J Pers Soc Psychol 54 (1988) 1063–1070.

[26] N.S. Jacobson William C Follette, D. Revenstorf, Psychotherapy Outcome Research: Methods for Reporting Variability and Evaluating Clinical Significance, Kazdin & Wilson, 1984.

[27] B.L. Jones, Proc Traj: A SAS Procedure for Group Based Modeling of Longitudinal Data, 2007. http://www.andrew.cmu.edu/~bjones.

[28] D.A. Kenny, Cross-lagged panel correlation: A test for spuriousness, Psychol Bull 82 (1975) 887–903. https://doi.org/10.1037/0033-2909.82.6.887.

[29] J. De Vries, H. Michielsen, G. Van Heck, Assessment of fatigue among working people: a comparison of six questionnaires, Occup Environ Med 60 (2003) 10–15. www.occenvmed.com.

Tables

Table 1. Questionnaire completion rates (number of participants that filled out this questionnaire at 2, 3, or 4 out of 4 measurement moments) by fatigue trajectory class.

	Total sample	Non-improvement class (23%)	Improvement class (9%)	Moderate improvement class (23%)	Weak improvement class (45%)
CIS-20	4: 121 3: 120 2: 52	4: 25 3: 27 2: 13	4: 13 3: 8 2: 3	4: 30 3: 23 2: 11	4: 53 3: 62 2: 25
PHQ-9	4: 125 3: 113 2: 57	4: 26 3: 24 2: 16	4: 13 3: 9 2: 2	4: 30 3: 22 2: 12	4: 56 3: 58 2: 27
PHQ-15	4: 125 3: 116 2: 55	4: 26 3: 24 2: 16	4: 13 3: 9 2: 2	4: 30 3: 23 2: 12	4: 56 3: 60 2: 25
GAD-7	4: 123 3: 119 2: 52	4: 26 3: 25 2: 15	4: 13 3: 9 2: 2	4: 28 3: 25 2: 11	4: 56 3: 60 2: 24
PANAS	4: 107 3: 121 2: 57	4: 22 3: 28 2: 10	4: 8 3: 13 2: 1	4: 29 3: 21 2: 13	4: 48 3: 59 2: 33
PSQ	4: 39 3:54 2:145	4: 7 3: 14 2: 31	4: 1 3: 2 2: 15	4: 14 3: 10 2: 28	4: 17 3: 28 2: 71
SF-36	4: 42 3:49 2:139	4: 8 3: 12 2: 30	4: 1 3: 2 2: 12	4: 15 3: 9 2: 27	4: 18 3: 26 2: 70

Note: CIS-20: Checklist Individual Strength. SF-36: MOS 36-Item Short-Form Health Survey. PHQ-15: Patient Health Questionnaire – 15 (somatic symptom severity). PHQ-9: Patient Health Questionnaire – 9 (depressive symptom severity). GAD-7: Generalized Anxiety Disorder – 7. PSQ: Perceived Stress Questionnaire. PANAS: Positive and Negative Affect Schedule.

Table 2. Average questionnaire scores of chronic fatigue syndrome patients following cognitive behavioral therapy before, during, and after treatment

	Before		**Follow-up 1**		**Follow-up 2**		**After**				
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**F**	**p**	**η2**
CIS-20											
Fatigue	51.4[a]	4.8	48.3[b]	8.2	46.6[c]	9.3	44.4[d]	11.	36.05	<0.001	0.08
Motivation	17.7[a]	5.8	17.5[a]	6.0	16.5[b]	6.0	16.0[b]	6.4	6.43	<0.001	0.02
Concentration	27.4[a]	6.0	25.5[b]	6.8	25.2[b]	7.1	23.9[c]	7.8	14.50	<0.001	0.04
Activity	16.0[a]	4.6	14.7[b]	5.3	14.2[bc]	5.0	13.8[c]	5.3	12.86	<0.001	0.03
Total	*112.5[a]*	*13.*	*104.0[b]*	*19.*	*102.6[c]*	*21.*	*98.1[d]*	*25.*	*25.85*	*< 0.001*	*0.06*
SF-36											
Physical	31.6[a]	8.3	31.3[a]	6.6	33.5[b]	7.7	35.3[b]	8.9	12.81	<0.001	0.03
Mental	37.7	11.	39.5	11.	39.00	11.	39.8	12.	1.20	0.31	0.00
PHQ-15	16.7[a]	4.7	16.8[b]	4.9	15.2[bc]	5.0	15.0[c]	5.7	12.00	<0.001	0.03
PHQ-9	12.8[a]	4.5	11.3[b]	5.2	10.5[c]	5.2	9.4[d]	5.6	25.08	< 0.001	0.06
GAD-7	11.6[a]	5.9	8.5[b]	4.7	7.4[c]	5.0	7.1[c]	5.0	41.38	<0.001	0.09
PSQ											
Perceived stress index	0.53[a]	0.1	0.51[ab]	0.1	0.49[bc]	0.1	0.46[c]	0.2	7.23	0.001	0.02
PANAS											
Positive affect	22.7	6.2	22.4	6.5	23.5	7.2	25.2	7.4	11.95	<0.001	0.03
Negative affect	22.9	8.0	23.2	8.0	21.9	8.0	21.3	7.3	4.85	0.003	0.01

Note: numbers in the same row with the same superscript are not significantly different from one another. CIS-20: Checklist Individual Strength. SF-36: MOS 36-Item Short-Form Health Survey. PHQ-15: Patient Health Questionnaire – 15 (somatic symptom severity). PHQ-9: Patient Health Questionnaire – 9 (depressive symptom severity). GAD-7: Generalized Anxiety Disorder – 7. PSQ: Perceived Stress Questionnaire. PANAS: Positive and Negative Affect Schedule.

Table 3. Intercept (initial fatigue score) and slope (change in fatigue score) of the different fatigue trajectories of chronic fatigue syndrome patients following cognitive behavioral therapy as estimated by latent class growth analysis.

Class	**Nr. of patients (%)**	**Intercept**		**Slope**		**Slope**
		Mean	**SE**	**Mean**	**SE**	**p-value**
Non-improvement	66 (23%)	123.98[c]	1.75	1.01[d]	0.89	0.25
Improvement	24 (9%)	107.68[ab]	2.51	-20.86[b]	1.45	< 0.001
Moderate improvement	66 (23%)	100.79[a]	2.34	-6.25[a]	1.09	< 0.001
Weak improvement	142 (45%)	112.39[b]	1.37	-3.16[c]	0.76	< 0.001

Note: numbers in the same column with the same superscript are not significantly different from one another. SE: standard error. Fatigue was quantified as the total score of the Checklist Individual Strength.

Table 4. Demographics and average questionnaire scores before treatment of the different fatigue trajectory classes of chronic fatigue syndrome patients following cognitive behavioral therapy as estimated by latent class growth analysis.

	Non-improvement class (23%)		Improvement class (9%)		Moderate improvement class (23%)		Weak improvement class (45%)		Statistic	p-value
% Women	87.9%		87.5%		84.8%		84.4%		$\chi^2(3)= 0.54$	0.91
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		
Age	44.8^{a}	8.5	39.8^{ab}	7.5	37.9^{b}	11.1	40.7^{b}	9.3	$F_{3,293} = 6.12$	< 0.001
Symptom duration	5.5	5.3	5.4	5.1	5.2	5.6	5.1	5.7	$F_{3,291} = 0.10$	0.96
CIS-20	124.8^{a}	9.0	113.6^{b}	14.0	101.1^{c}	12.9	112.0^{b}	10.7	$F_{3,280} = 47.4$	< 0.001
SF-36 physical	30.9	9.3	31.9	10.3	33.4	8.2	30.8	7.5	$F_{3,249} = 1.54$	0.20
SF-36 mental	34.0^{a}	11.7	36.0^{ab}	13.6	40.0^{b}	10.2	38.5^{ab}	11.6	$F_{3,249} = 2.95$	0.03
PHQ-15	18.0	4.9	16.4	5.6	16.2	4.4	16.4	4.6	$F_{3,290} = 1.90$	0.13
PHQ-9	15.4^{a}	4.9	12.1^{b}	5.1	11.4^{b}	4.2	12.4^{b}	4.4	$F_{3,289} = 9.82$	< 0.001
GAD-7	13.3	5.3	11.2	6.5	10.7	5.8	11.2	6.0	$F_{3,290} = 2.67$	0.048
PSQ index	0.60^{a}	0.19	0.52^{ab}	0.18	0.47^{b}	0.17	0.52^{b}	0.16	$F_{3,261} = 5.37$	0.001
PANAS – positive	20.0^{a}	6.2	23.5^{abc}	8.0	25.6^{b}	6.0	22.6^{c}	5.6	$F_{3,256} = 8.74$	< 0.001

Note: numbers in the same row with the same superscript are not significantly different from one another. CIS-20: Checklist Individual Strength. SF-36: MOS 36-Item Short-Form Health Survey. PHQ-15: Patient Health Questionnaire – 15 (somatic symptom severity). PHQ-9: Patient Health Questionnaire – 9 (depressive symptom severity). GAD-7: Generalized Anxiety Disorder – 7. PSQ: Perceived Stress Questionnaire. PANAS: Positive and Negative Affect Schedule

Table 5. Risk factors for being allocated to the improvement, moderate improvement, and weak improvement classes relative to the non-improvement class for fatigue trajectories of chronic fatigue syndrome patients following cognitive behavioral therapy.

	Improvement			**Moderate improvement**			**Weak improvement**		
	β	**SE**	**p**	**β**	**SE**	**p**	**β**	**SE**	**p**
Depressive symptoms	-0.11	0.06	0.057	-0.27	0.07	<0.001	-0.16	0.04	<0.001
Somatic symptom severity	-1.6	1820	<0.99	-0.13	0.05	0.009	-0.11	0.04	0.004
Anxiety symptoms	-0.07	0.05	0.14	-0.11	0.04	0.012	-0.07	0.03	0.029
Positive affect	0.17	0.05	<0.001	-1.24	3679	<0.99	0.10	0.04	0.005
Perceived stress	-2.71	1.79	.13	-5.56	1.54	<0.001	-3.32	1.24	0.008

Note. Depressive symptoms: Patient Health Questionnaire – 9 score. Somatic symptom severity : Patient Health Questionnaire – 15 score. Anxiety symptoms: Generalized Anxiety Disorder – 7 score. Positive affect: Positive affect subscale of the Positive and Negative Affect Schedule. Perceived Stress: perceived stress index derived from the Perceived Stress Questionnaire. For the improvement and moderate improvement classes an extremely large standard error was found for positive affect and somatic symptom severity respectively, indicative of an irregularity in the model.