

An analysis of Dutch hallmark studies confirms the outcome of the PACE trial: cognitive behaviour therapy with a graded activity protocol is not effective for chronic fatigue syndrome and Myalgic Encephalomyelitis

FNM Twisk^{1*} and LAMM Corsius²

¹ME-de-patiënten Foundation, Limmen, the Netherlands

²Independent researcher, Etten-Leur, The Netherlands

Abstract

Myalgic Encephalomyelitis (ME) and Chronic Fatigue Syndrome (CFS) are considered to be enigmatic diseases. Several studies propose that the combination of cognitive behaviour therapy with a graded activity protocol (CBT+), justified by a so-called (bio)psychosocial (explanatory) model, is an effective treatment option for CFS (ME).

Objective A critical review of five Dutch hallmark studies that allegedly support this claim.

Methods An analysis of the five CBT+ studies with special attention to the patients studied, the criteria (subjective and objective measures and cut-off scores) used to select participants and to define improvement and recovery, the consistency of the definitions of caseness (being diagnosed as a CFS patient at entry) versus the definitions of improvement and recovery after CBT+, and the objective effects.

Results The studies investigated suffer from various methodological flaws. Apart from these methodological shortcomings, the claim that CBT+ is an effective treatment option for CFS is not substantiated by the data reported. Some studies investigated CFS patients, other studies investigated CF patients, labelled as CFS patients, or combinations of CFS and CF patients. No study investigated the effect of CBT+ in a group of patients meeting the (original) diagnostic criteria for ME. The effects of CBT+ on subjective measures, for example fatigue and disability, if present, are insufficient to achieve normal values. Impressive recovery and improvement rates are based on very loose criteria for subjective measures. Cut-off scores for subjective measures used to define improvement and recovery in studies show overlap with cut-off scores for CFS caseness in one or more of the other studies. More importantly, looking at the objective measures, the proof of clinical improvement after CBT+ is lacking.

Conclusion Solid evidence of effectiveness of CBT+ for CFS, let alone ME, is lacking in the five hallmark studies. The lack of objective improvement indicates CBT+ is ineffective. This finding confirms the outcome of the large-scale PACE-trial in the UK.

Introduction

Myalgic Encephalomyelitis (ME) [1-3] and Chronic Fatigue Syndrome (CFS) [4] are considered to be controversial diseases [5,6]. The dispute largely originates from two opposing paradigms for the etiology and the therapies for ME and CFS: the biomedical and the (bio)psychosocial model. The (bio)psychosocial model assumes a clear distinction between initiating factors, e.g. infections, predisposing factors, e.g. stress, and 'illness-perpetuating' factors [7]. In this model all symptoms can be fully explained by psychosocial 'illness-perpetuating' factors (cognitions and behaviour), which are fully independent from the initiating factors. Justified by this model two types of interventions have been developed: CBT, targeting "cognitive responses (fear of engaging in activity) and behavioural responses (avoidance of activity), the latter being responsible for the symptoms", and GET or graded activity (GA) aimed at gradually increasing activity to reverse 'deconditioning' [8].

The aim of this study was to investigate if the claim that the combination CBT/GA (CBT+) is an effective intervention for CFS

and ME patients and the claim that CBT+ is a safe intervention are substantiated by the outcomes of five studies in the Netherlands. In order to position the analysis of the five studies and the overall analysis, it is relevant to discuss two factors directly related to the claim that CBT+ is effective for CFS (ME): the participants studied (selection criteria: diagnostic criteria and additional criteria), and the type of measures to assess the symptoms (in order to diagnose patients) and to determine the effectiveness of interventions.

Correspondence to: Frank NM Twisk MBA MBI BEd BEc, ME-de-patiënten Foundation, Limmen, The Netherlands, Tel: +31-72-505 4775; E-mail: frank.twisk@hetnet.nl

Key words: Myalgic Encephalomyelitis, chronic fatigue syndrome, cognitive behaviour therapy, graded exercise therapy, measures, effectiveness

Received: November 02, 2017; **Accepted:** November 06, 2017; **Published:** November 10, 2017

Diagnostic criteria for ME, CFS and chronic fatigue: different patient populations

One relevant factor in the assessment of studies into interventions is the patient selection criteria: diagnostic criteria and additional criteria, e.g. fatigue scores. This paragraph outlines the relevant diagnostic criteria for patient selection (eligibility). ME [1-3] is a neuromuscular disease with “a marked similarity to non-paralytic poliomyelitis in respect of prodrome, seasonal and geographical incidence” [3]. ME [1-3], (often) initiated by an infection, is characterized by a) distinctive muscular symptoms, including myalgia, muscle tenderness, and muscle weakness lasting for days after trivial exertion; b) neurological symptoms associated with cognitive, autonomic and sensory functions; e.g. concentration and memory deficits, sleep reversal and emotional lability, and c) variable involvement of cardiac and other systems, e.g. cold extremities, loss of thermostatic stability, and orthostatic tachycardia. ME is accompanied by “an unpredictable state of central nervous system exhaustion following mental or physical exertion which may be delayed and require several days for recovery” [9], labeled post-exertional neuro-immune exhaustion by the International Consensus Criteria for ME (ME/ICC) [10].

The only mandatory feature of CFS is chronic fatigue. To meet the diagnosis CFS [4] (medically unexplained, incapacitating) fatigue must be accompanied by at least four out of the following eight symptoms: impairment in short-term memory or concentration; sore throat; tender lymph nodes; muscle pain; multi-joint pain without swelling or redness; headaches of a new type, pattern, or severity; unrefreshing sleep; and post-exertional ‘malaise’ lasting more than 24 hours. CFS [4] is not equivalent to (a severe form of) ME [1-3,11], although patients can meet both diagnoses. That’s not a matter of opinion [12], but a matter of definition [5]. A patient can meet the diagnosis CFS [4], while not experiencing any of the distinctive ME symptoms [1-3], and ME patients can fail to meet the diagnosis CFS [13].

To complicate the diagnostic issue even further, some authors interpret CFS to be equivalent to chronic fatigue (CF) (without any of the eight additional symptoms) [7,14]. When evaluating studies into the effects of CBT and GET or GA, it is crucial to establish the patients studied: patients with ME, CFS or (a specific variant of) CF.

Assessment of the effects of CBT, GET and GA: subjective versus objective measures

Another relevant factor in the evaluation of studies into interventions is the type of measures used to assess/diagnose patients at baseline and after intervention. Patient selection and the assessment of the effects of interventions (CBT, GET, et cetera), including definitions of clinical improvement and recovery, are often based on non-specific, subjective measures, e.g. fatigue [15,16], physical functioning [17] or disability [18]. Studies into (behavioural) interventions less frequently use objective measures, e.g. physical exercise capacity, activity levels or re-employment. This observation is relevant, because subjective measures are associated with a placebo response [19], response bias [20], researcher allegiance [21], et cetera. Moreover, when considering ME/CFS, subjective measures don’t seem to correlate with objective measures. For example, one study [22] found that fatigue was not correlated with maximum oxygen uptake, another study [23] observed that self-reported cognitive impairment is not related to cognitive test scores. To assess the effects of interventions impartially the use of objective measures is essential [24].

Methods

To investigate if the proposed effectiveness of CBT+ in ME and CFS is substantiated by the data, we carried out an analysis of five hallmark studies conducted in the Netherlands, with a focus on the patients investigated in the studies, the subjective and objective measures and cut-off scores used to select participants (to define caseness) and to define improvement or recovery, the consistency of the definitions of caseness and definitions of improvement or recovery, and the objective effects. The five studies included in this analysis were selected for several reasons: the studies have been cited very frequently, have attracted a lot of attention by the media and other researchers, were conducted in the Netherlands by one group of researchers, have investigated various variants of CBT+ (face-to-face CBT+: individual and group sessions, CBT+ for adolescents, and CBT+ by the internet), and have a strong impact on the medical policies with regard to ME and CFS in the Netherlands [25-27]. It is important to note that while the authors label the intervention ‘CBT’ throughout their studies, the intervention investigated in all five studies is a combination of CBT and GA [7,28]. For that reason, the intervention is labelled as CBT+ in this article. The five studies analysed in this review are: a) the Prins *et al.*, 2001 trial (trial [7] and related outcome studies [23,29]); b) the Bazelmans *et al.*, 2005 trial [30]; c) the Stulemeijer *et al.*, 2005 trial (trial [31], related outcome studies: [23,33] and follow-up study [34]); d) the Knoop *et al.*, 2007 study [35]; and e) the Nijhof *et al.*, 2012 (Fatigue In Teenagers on the interNET – FITNET) trial (trial [36], protocol study [37], long-term follow-up study [38], commentaries [39,40], and editorials [41,42]). During our analysis we found that the studies suffer from various methodological flaws, partially inherently related to the type of intervention, e.g. design of the control condition and the impossibility to blind the participants. Methodological issues will be briefly addressed to in the Discussion section. This analysis however focuses on the question if the claims with regard to the effectiveness and safety of CBT+ for CFS and ME are substantiated by the data reported.

Results

The characteristics of the studies analysed are described in Table 1. The subjective and objective measures used in the five studies are explained in Table 2. As can be seen in Table 1, there are significant differences in the studies with regard to the patients studied, the type of intervention employed, and the measures and cut-off scores used to define caseness (entry criteria), improvement and recovery. In the next paragraphs the studies and their outcomes will be discussed in more detail.

Prins *et al.* (2001)

Prins *et al.*, 2001 [7] concluded that “[CBT+] was significantly more effective than both control conditions for fatigue severity [...] and for functional impairment [...]” in CFS at 14 months. Both CBT+ and non-intervention had a positive effect on the means of the two primary outcomes, the CIS fatigue (CIS F) subscore [16] and the SickNES Impact Profile (SIP 8) score [18], although the positive effect of CBT+ on these measures was significantly larger. But the effects of CBT+ (and non-intervention) were by far insufficient to achieve ‘normal levels’ as defined in this study and other studies by the same research group for CIS F [30,35] and SIP 8 [7, 35]. At the group level there was no significant difference between the effect of CBT+ and non-intervention on the secondary subjective measures, except for the Karnofsky status [43,44], which was rated by a clinical psychologist, not by the patient. While the study [7] states that “The final goal of [CBT+] for CFS included work rehabilitation”, employment rates before and after CBT+ and non-

Table 1. Characteristics of the five hallmark studies analysed

| Study | Prins et al., 2001 [7] | Bazelmans et al., 2005 [30] | Stulemeijer et al., 2005 [31] | Knoop et al., 2007 [35] *a | Nijhof et al., 2012 [36] |
|--|--|---|---|---|---|
| Secondary outcome studies | [23,29] | | [23,29,32,33] | | |
| Protocol study | | | | | [37] |
| Follow-up study | | | [34] | | [38] |
| Commentaries | | | | | [39,40] |
| Editorials | | | | | [41,42] |
| Participants | | | | | |
| Diagnostic criteria | Chronic fatigue *b | CFS (Fukuda et al., 1994) or idiopathic chronic fatigue *c | CFS (Fukuda et al., 1994) | CFS (Fukuda et al., 1994), with the exception that comorbidities that could explain 'fatigue' were not an exclusion criterion | CFS (Fukuda et al., 1994) |
| Inclusion criteria | CIS F ≥ 40 ("severe fatigue"), and SIP ≥ 800 ("severe functional impairment") | CIS F ≥ 35 , SIP ≥ 700 , and willing to stop other treatments during the trial | CIS F ≥ 40 ("severe fatigue"), and SF-36 PF ≤ 65 ("severe functional impairment") and adolescent | CIS F ≥ 35 ("severely fatigued"), and SIP ≥ 700 ("severely disabled") | CIS F ≥ 40 ("severe fatigue") and CHQ-CF87 ≤ 85 or school attendance *d $\leq 85\%$ ("functional impairment"), and adolescent |
| Eligible patients refusing to participate in the trial | 99/377 (26.3%) | 62/129 (48.1%): 19 patients didn't want to stop other treatments, and 43 patients refused for other reasons | 11/82 (13.4%) | ? | 6/141 (3.5%): 3 patients refused to participate, and 3 patients preferred face-to-face CBT |
| N | 241 *e | 67 | 69 *f | 96 *g | 135 |
| Mean age CBT+*h arm | 36.2 | 37.4 | 15.6 | 37.0 | 15.9 |
| Intervention | | | | | |
| Description | CBT+ *h individual face-to-face sessions | CBT+ *h face-to-face in groups (CBGT+) | CBT+ *h individual face-to-face sessions, two variants of CBT+ depending on the type of activity pattern of the patient | CBT+ *h individual face-to-face sessions | CBT+ *h individual sessions through the internet (FITNET) |
| N | 83 | 31 | 35 (pervasively passive: 10, relatively active: 23, unknown: 2) | 16 | 68 |
| Control condition | | | | | |
| Description | Natural course (nonintervention) (NC) | Waiting list (non-intervention) (WL) | Waiting list (non-intervention) (WL) | *a | 'Usual care' (UC): CBT+ *h (65.6%), GET (49.3%), 'rehabilitation therapy' (49.3%) and/or 'alternative therapies' (23.9%) or no therapy at all (10.4%) |
| N | 78 | 36 | 34 | | 67 |
| | Guided support in groups (GSG) | | | | |
| N | 80 | | | | |
| Drop-out | | | | | |
| Intervention group | 1/93 (0.1%) was excluded during the trial, 10/93 (10.8%) didn't start CBT+ *h, and 23/82 (29.0%) 'dropped out' | 2/31 (6.5%) dropped out | 1/36 was excluded after randomisation, and 6/35 (17.1%) dropped out: 3 did not start and 3 withdrew | 3/112 (2.7%) did not start, and 13/109 (11.9%) dropped out during the trial | 1/68 (1.4%) |
| Control group(s) | NC: 3/91 (3.3%) was excluded, 9/88 (10.2%) withdrew from the trial | | 1/35 (2.8%) was excluded during the trial, and 1/34 (2.9%) 'dropped out' | *a | 3/67 (4.4%) |
| | GSG: 4/94 (4.2%) was excluded during the trial, 8/90 (8.8%) didn't start GSG, and 17/90 (18.8%) dropped out | | | | |
| Complete data | | | | | |
| | At 8 months: CBT+ *h: 83 GSG: 80, and NC: 78 | Number of patients with complete data at 6 months were not reported. | At 5 months: CBT+ *h: 29 and WL: 33 | After intervention: CBT+ *h: 96 | At 6 months: FITNET: 67, and UC: 64 |
| | At 14 months: CBT+ *h: 58, GSG: 62, and NC: 76 | SRI scores were available from 27 of 29 patients in the CBGT+ arm. SRI scores from the WL arm not reported. | | | At 12 months: FITNET: 64, and UC: 63 |

| Study | Prins et al., 2001 | Bazelmans et al., 2005 | Stulemeijer et al., 2005 | Knoop et al., 2007b *a | Nijhof et al., 2012 |
|---|--|---|--|---|---|
| Outcomes *i | | | | | |
| Primary outcome measures | | | | | |
| | At the start, at 8 months (after intervention) and at 14 months | At the start and at 6 months (after intervention) | At the start and after intervention (at 5 months) | At the start and after intervention | At the start, at 6 months (after intervention), and at 12 months |
| Subjective measures | CIS F and SIP 8 | CIS F and SIP 8 | CIS F and SF-36 PF | CIS F and SIP-8 *j | CIS F and CHQ-CF87 |
| Objective measures | | | School attendance *d | | School attendance *d |
| Secondary measures | | | | | |
| Subjective measures | KPS (rated by a clinical psychologist), SCL-90, EQL-5D VS and self-rated improvement (SRI): "completely recovered", "much better", unchanged, or "worse" | Daily observed fatigue (DOF), daily observed pain (DOP), SCL-90, BDI, and SRI: "completely recovered", "better or much better", "the same", or "worse" *k | Severity of the nine symptoms defining CFS (Fukuda et al., 1994) using a 4 Likert scale, ranging from "never" (1) to "every day" (4) | SF-36 PE, SF-36 SF, SF-36 GH, and number of factors of the FQL scoring negative | SRI: "completely recovered", "much better", "the same complaints" of "become worse" |
| Objective measures | Hours worked in a job in 12 days, and physical activity levels (assessed by an actometer) *l | Hours worked in the last week | Physical activity levels (assessed by an actometer) *l | | Physical activity levels (assessed by an actometer) *m |
| Study | Prins et al., 2001 | Bazelmans et al., 2005 | Stulemeijer et al., 2005 | Knoop et al., 2007b *a | Nijhof et al., 2012 |
| Criteria for improvement or recovery at the individual level | Improvement | Improvement | Improvement | Recovery | Recovery |
| | Clinically significant improvement in fatigue severity: CIS F ≤ 36 (reliable change index >1.64 , $p < 0.05$) | | Clinically significant improvement in fatigue severity: CIS F < 35.7 (and reliable change index >1.96) | Recovery from CFS: CIS F < 35 and SIP 8 < 700 | |
| | Clinically significant improvement in KPS: an improvement of ≥ 10 points and a KPS score ≤ 80 | Improvement: SRI is "completely recovered" or "much better" | Clinically significant improvement in functional impairment: change in SF-36 PF ≥ 50 or SF36 PF score ≥ 75 | Most comprehensive definition of recovery *n: CIS F ≤ 27 ("level of fatigue comparable to healthy people"), SF-36 PF ≥ 80 ("no physical disability"), SF-36 SF ≥ 75 ("no social disability"), SF-36 GH ≥ 65 ("normal health perception"), and number of factors of the FQL scoring negative = 0 ("no negative perception of fatigue") | CIS F < 40 , CHQ-CF87 ≥ 85 , school attendance $\geq 90\%$, and SRI: "completely recovered" or "much better but still experiencing some symptoms" |
| | Clinically significant self-rated improvement: "completely recovered" or "much better" | | Clinically significant improvement in school attendance: fully attending school | | |
| | | | Clinically significant self-rated improvement: "completely recovered" or "much better but still experiencing some symptoms" | | |

*a The Knoop et al., 2007b study was not a randomized controlled trial, since a control group was lacking.

*b Fukuda criteria for CFS with "the exception of the criterion requiring 4 of 8 additional symptoms to be present".

*c According to the study 2 participants experienced idiopathic chronic fatigue and did not meet the Fukuda criteria for CFS.

*d School attendance: attended lessons divided by total lessons to attend in the previous week.

*e Number of patients with complete data at 8 months.

*f In both groups one patient was excluded after randomisation because they didn't met the CFS criteria.

*g 22 of the 96 patients (23%) reported one or more comorbidities that could explain 'chronic fatigue'. A clinical condition that could explain 'chronic fatigue' is an exclusionary criterion for CFS.

*h The authors refer to the intervention as 'CBT' throughout the studies.

However, the studies employed protocols in which CBT is combined with a gradual increase of activity levels (hence: CBT+).

*i This table summarizes clinical measures only.

Measures assessing satisfaction with therapy et cetera and variables assessing supposed perpetuating factors are excluded.

*j The SIP 8 score was not included in "the most comprehensive definition of recovery".

*k Self-rated improvement scores were only reported for the CBGT group (27 patients, data missing from 2 patients).

*l Physical activity levels (assessed with an actometer) were not reported until 2010 (Wiborg et al., 2010).

*m Physical activity levels (assessed with an actometer) as described in the protocol (Nijhof et al., 2011) were not reported.

*n The study also reports recovery on specific criteria, e.g. fatigue (CIS F ≤ 27) and "disabilities in all domains" (SIP 8 ≤ 203).

Table 2. Measures used in five Dutch studies into the effects of CBT+.

| Measure | Abbreviation | Number of items | Possible scores for one item | Range | +/- *a |
|--|--------------|------------------------|--|----------------|--------|
| Beck Depression Inventory | BDI | 21 | 0-3 | 0-63 | - |
| Child Health Questionnaire – Child Form / Physical Functioning subscale | CHQ-CF87 PF | 3 | | 0-100% | + |
| Checklist Individual Strength / Fatigue subscale | CIS F | 8 | 1-7 | 8-56 | - |
| Daily Observed Fatigue: mean of the total daily observed fatigue scores four times a day over 12 days | DOF | 4 | from “not fatigued” (0) to “very severely fatigued” (4) | 0-16 | - |
| Daily Observed Pain: mean of the total daily observed pain scores four times a day over 12 days | DOP | 4 | from “no pain” (0) to “very severe pain” (4) | 0-16 | - |
| European Quality of Life Index / Visual Analogue scale | EQL-5D VS | 1 | 1 | 0-100 | + |
| Fatigue Quality List, consisting of 18 adjectives describing the experience of fatigue, yielding 4 factors, 3 of which have negative connotations of fatigue: “frustrating”, “exhausting” and “frightening”. | FQL | 3 | “experience not applicable” (0) or “experience applicable” (1) | 0-3 | - |
| Karnofsky Performance Status | KPS | 1 | from “dead” (0) to “normal no complaints; no evidence of disease” (100) | 0, 10, ...,100 | + |
| Symptom Checklist: 90 items | SCL-90 | 90 | 1-5 | 90-450 | - |
| Medical Outcomes Study (MOS) 36 Item Short Form Health Survey (SF-36), consisting of two dimensions: the physical health dimension (21 items), consisting of the 1. physical functioning (10 items), 2. role limitations due to physical health (4 items), 3. bodily pain (2 items), and 4. general health (5 items), subscales and the mental health dimension (14 items), consisting of 5. vitality/energy (4 items), 6. role limitations due to emotional problems (3 items), 7. emotional well-being (5 items), and 8. social functioning (2 items) subscales, each subscale yielding a score of 0-100, dividing the total score by the number of items | SF-36 | 8 subscales (36 items) | 0-100 | 0-800 | + |
| SF-36 / General Health subscale (weighted average of the score on the relevant items) | SF-36 GH | 5 | 0, 25, 50, 75 or 100 | 0-100 | + |
| SF-36 / Physical Functioning subscale (weighted average of the score on the relevant items) | SF-36 PF | 10 | “Yes, limited a lot (0), “yes, limited a little” (50), or “no, not limited at all” (100) | 0-100 | + |
| Medical Outcomes Study (MOS) 36 Item Short Form Health Survey (SF-36) / Social Functioning subscale (weighted average of the score on the relevant items) | SF-36 SF | 2 | 0, 25, 50, 75 or 100 | 0-100 | + |
| Sickness Impact Profile-8: total weighted score of 8 of the 12 subscales: home management (10 items), mobility (10 items), alertness behaviour (10 items), sleep/rest (7 items), ambulation (12 items), social interaction (20 items), work (9 items) and recreation/pastimes (8 items) | SIP 8 | 86 | | 0-5.799 | - |

*a + higher scores implicate better health, - lower scores indicate better health

intervention were not reported. However, there were no significant differences between the effects of CBT+ and non-intervention on the number of hours working in a job, while a secondary outcome study [29] showed that the non-significant, extremely modest effect of CBT+ on activity levels was by far insufficient to achieve normal levels. Another secondary outcome study [23] found that CBT+ has no significant effect on a third objective measure, cognitive test scores. All in all, the effects on CBT+ at the group level are non-existent or marginal and insufficient to achieve normal levels. On the individual level the study reported that clinically significant improvement was seen in CIS F for 20 of 58 (35%), in Karnofsky performance status for 28 of 57 (49%), and self-rated improvement for 29 of 58 (50%).” But the criteria used for the definition of clinical improvement on these three measures seem very loose, since 32% of the patients in the waiting-list group also reported a clinical significant improvement without any intervention and 23% experienced a significant improvement of the Karnofsky performance status. Oddly enough, clinically significant improvement of the SIP 8 score, one of the two criteria to define caseness

and to include patients in the trial, was not reported. Looking at the effects of CBT+ and non-intervention at the group level (means), not many patients, if any, would achieve a clinical significant improvement on this measure. So, CBT+ yielded modest effects on some measures and no effects on most measures, especially three objective measures, including number of hours worked.

Most importantly, the trial by Prins *et al.* [7] didn’t select CFS [4] patients, as stated, but patients suffering from CF. In addition, as the authors state “There was a large withdrawal rate [...], especially in the CBT and support groups.” [7]. In the CBT+ group 10/93 (10.8%) didn’t start and 23/83 (27.7%) dropped out during the trial. The dropout/withdrawal rate is very high considering the fact that patients with severe symptoms couldn’t participate and the trial most likely suffered from a high rate of self-selection beforehand, due to the fact that CFS patients seem to be skeptical of psychological interventions, like CBT+ [35]. The latter is illustrated by the finding that 99 (20.8%) of the 476 eligible patients refused to participate in the trial.

Bazelmans *et al.*, 2005

In a trial by Bazelmans *et al.*, 2005 [30], both CBT+ and non-intervention had a very modest effect on mean CIS F fatigue scores after 6 months. Although the effect of CBT+ (CIS F at baseline: 51.0, at 6 months: 45.6) was significantly greater than the effect of non-intervention (CIS F at baseline: 50.8, at 6 months: 48.4), the treatment effect (3.0 in comparison with the non-intervention group) was largely insufficient to achieve normal levels as defined by another study of the research group (≤ 27 [35]). The mean CIS F score after CBT+ would be qualified as “severe fatigue” according to criteria of other studies by the same research group (≥ 40 [7], ≥ 35 [45]) and indicates that patients were still ill enough to re-enter the trial for which a score of ≥ 35 was required. Curiously non-intervention had a positive effect on functional impairment (mean SIP 8 score at baseline 1,710, at 6 months: 1,417), but CBT+ had a small negative effect on the mean SIP 8 score (before: 1,707, after: 1,736). As the authors phrased it: “For functional impairment, the effect was opposite to what was expected”. Looking at the improvement of SIP 8 in the non-intervention group one could argue that CBT+ is impeding the naturally-occurring recovery process. In both groups the mean SIP 8 scores after the trial were sufficient to label patients as ‘severely disabled’ (≥ 800 [7], ≥ 700 [35,46]) and were largely insufficient to reach ‘normal levels’ (≤ 203 [35]). CBT+ and non-intervention had no effect on the four secondary subjective measures (daily observed fatigue, daily observed pain, psychological well-being and depression). Most importantly, since work resumption is the final goal of CBT+ [7,35], CBT+ had no effect on the (very low) number of hours worked. At the individual level 37% of the patients in the CBT+ arm rated themselves as “completely recovered” or “(much) better”. Despite this, CIS fatigue and SIP 8 scores remained high in the CBT+ group. Self-rated improvement was not assessed in the waiting list group. However, in another study [7] 32% of the patients in the non-intervention arm reported a clinical significant improvement afterwards. Most patients who ‘improved’ or ‘recovered’ by CBT+ where patients with less fatigue and less pain and significantly less disability (lower SIP 8 scores) at baseline.

Stulemeijer *et al.*, 2005

Stulemeijer and others [31] studied CBT+ in adolescents and concluded that “Patients in the [CBT+] group reported significantly greater decrease in fatigue severity [...] and functional impairment and their attendance at school increased significantly [...]”, that “They also reported a significant reduction in several accompanying symptoms.”, and that “Self-reported improvement was largest in the study group.”. Drop-out rates were significant. A substantial subgroup of patients didn’t start or withdrew from CBT+: “Six patients (19%) withdrew from therapy [...]” [31]. Both CBT+ and non-intervention had a positive effect on fatigue (CIS F score), physical functioning (MOS 36-item short-form health survey SF-36 PF score [17]), although the effects of CBT+ on these two measures were significantly higher. While the effect of CBT+ was substantially larger, both CBT+ and non-intervention had a positive effect on the only objective measure, school attendance. Despite this, school absence in both groups remained rather high in both groups. The effect of CBT+ on the other eight CFS symptoms was very modest or non-existent. At the individual level, both CBT+ and non-intervention showed positive effects in substantial subgroups. No less than 44% in the non-intervention group rated themselves as “completely recovered” or “much better” (versus 71% in the CBT arm). Looking at the natural course of the disease [47] and recovery rates (without therapy) [48,49], one can question the value and relevance of

this measure (in this study and other studies) and/or the diagnosis CFS [4]. The same applies to school attendance, since 29% of the patients in the non-intervention group reported full school attendance at 5 months (versus 58% in the CBT+ group), fatigue severity (21% of patients in the non-intervention reported improvement: CIS F < 35.7 and a reliable change index > 1.96 vs. 60% in the CBT+ arm), and physical functioning (non-intervention group: 24%, CBT+ arm: 63%). The negligible effects of CBT+ on activity levels observed in a secondary outcome study [29] are at odds with less school absence after CBT+. Another secondary study [23] found that CBT+ didn’t yield an improvement of cognitive test scores. This latter observation is relevant, since a study by the same research group [50] found that CFS has a great impact on cognitive functioning. Although 15,3% of the adolescent with CFS had already switched to a lower school level, the IQ of CFS patients was still 8 points below their peers. More important, the decline in IQ was not due to school absence. According to a study by others [51] (reduced) cognitive functioning is not correlated to (higher) levels of fatigue. All in all, the data of the trial [31] and secondary outcome studies [23,29] do not support the conclusion that “[CBT+] is an effective treatment for CFS in adolescents.” [31].

Knoop *et al.*, 2007

Knoop *et al.*, 2007 [35] reported impressive recovery rates for CFS by CBT+: “After treatment, 69% of the patients no longer met the CDC criteria for CFS.”. But as the authors also acknowledge: “The percentage of recovered patients depended on the criteria used for recovery.”. According to Knoop *et al.*, 2007 [35] 23% of the CFS patients fully recovered using “the most comprehensive definition of recovery”. First of all, looking at the co-morbidities reported, one could question the correct application of the diagnostic criteria for CFS [4] to select patients in this study [35]. The diagnosis CFS [4] is only applicable when the patient doesn’t experience medical and psychological comorbidities which can adequately explain “fatigue”. The criteria for ‘recovery from CFS (CDC)’ are very easily met, e.g. an improvement from 35 to 34 for CIS F (range 8-46) combined with an improvement from 700 to 699 for SIP 8 (range 0-5.799) are sufficient to be qualified as being ‘recovered from CFS’ (CIS F < 35 and SIP8 < 700) in this study. However, these scores are by far insufficient to achieve the ‘normal levels’ as defined by the same study (CIS F ≤ 27 and SIP8 ≤ 203). This is illustrated by the observation that the positive effect of CBT+ on the mean SIP 8 score is by far insufficient to reach ‘normal levels’ defined in this study (≤ 203). Not surprisingly, the criteria employed to define recovery largely determine the ‘recovery rates’. Using ‘more strict’ criteria for recovery (CIS F ≤ 27 , SIP8 ≤ 203 , SF-36 SF-36 PF ≥ 80 , SF-36 Social Functioning subscore ≥ 75 , SF-36 General Health subscore ≥ 65 , and no factors scoring negative on the Fatigue Quality List), the recovery rate drops to 23%. However, even “the most comprehensive definition of recovery” isn’t based on stringent criteria. Curiously, the SIP 8 score, used as a criterion to select patients (caseness) in this study [35] and other trials [7], isn’t included in these two definitions of recovery. The study doesn’t report how many CFS patients reached ‘normal levels’ (≤ 203), but considering the size of the effect of CBT+ on SIP 8 in this study and other trials, few patients, if any, would reach ‘normal levels’. An important point of criticism on the ‘normal values’ used in this study and other studies by the research group relates to the method by which these ‘normal values’ are determined. The ‘threshold scores’ are defined as the mean ± 1 SD of the healthy population. However, as the authors acknowledge the SIP 8 and SF-36 PF are not normally distributed but skewed [35]. The same applies to CIS F [16] and other SF 36 subscales [52,53]. Aaronson *et al.*, 1998 [54], cited in

Knoop *et al.*, 2007 [35], showed a large ceiling effect of SF-36 PF: 31.9% of the Dutch population scored at the highest scale level. As Knoop *et al.* [35] state “Therefore one could argue that recovery according to the SIP8 has to be defined as scoring the same or lower than the 85th percentile of the healthy reference group.” Using percentiles as threshold scores instead of the mean \pm 1 SD for all subjective measures has a non-negligible negative effect on ‘recovery rates’: “[T]he recovery rate using the definition of having no disabilities in all domains [...] would decrease from 26 to 20%.”. Likely due to the use of the mean \pm 1 SD algorithm for calculating ‘normal values’ and/or the use of non-representative reference populations, the ‘normal value’ for SF-36 PF for the ‘young’ CFS patient group (mean age: 37.0 years) comes close to the mean SF-36 PF scores for healthy population of seniors aged 55 to 64 years [55], while the ‘normal values’ for SF-36 Social Functioning score resembles the mean score of older people aged 75 to 84 years [55] and the ‘threshold score’ for the SF-36 General health score is comparable with the mean of a population of 65 to 74 years [55]. If percentiles of representative populations were employed to define ‘normal values’ and the SIP 8 score was included in the “the most comprehensive definition of recovery” the ‘recovery rate’ based on the subjective measures used would drop dramatically.

The impressive recovery rates reported by the Knoop *et al.*, 2007 study [35] aren’t justified by the data, since the study lacked a control group and non-intervention showed to have positive effects on the subjective measures in substantial patient subgroups in other studies [7-31]. Furthermore, the effect of CBT+ on the other symptoms defining CFS [4] aren’t reported. The study lacked objective measures to substantiate ‘recovery’. Finally, the study reported much lower recovery rates for patients with comorbidities, while many CFS patients experience comorbidities [56].

Nijhof *et al.*, 2012

Nijhof *et al.*, 2012 [36,37] compared CBT+ delivered by the internet (FITNET: Fatigue In Teenagers on the interNET) in adolescents aged 12-18 years with ‘usual care’: CBT+ (66% of the patients in the ‘usual care’ group), physical treatment, in most cases GET (49%), rehabilitation treatment (22%), alternative treatment (24%), and no treatment (10%). Primary outcomes were fatigue (CIS F), physical functioning (CHQ PF: Child Health Questionnaire [57,58] Physical Functioning score) and school attendance at 6 months. FITNET was reported to have a substantial positive effect on the primary measures (CIS F, CHQ PF and school absence), while ‘usual care’ had smaller effects. The effect of FITNET and ‘usual care’ on the other eight symptoms defining CFS [4] were not reported. A follow-up study [38] showed that there were large differences between ‘recovered’ and ‘non-recovered’ patients and that the positive effects of FITNET on the mean scores were determined by the scores of ‘recovered patients’ to a large degree. Nijhof and colleagues reported extraordinary results: “FITNET was significantly more effective than was usual care for all dichotomised primary outcomes at 6 months - full school attendance (75% vs 16%), absence of severe fatigue (85% vs 27%), and normal physical functioning (78% vs 20%)” [36]. But an analysis of the data raises scepticism about the impressive results of FITNET. The criteria to select patients (“severe fatigue”: CIS F \geq 40, and functional impairment: CHQ PF \leq 85, or a school attendance \leq 85%) implicate that someone could already meet one of the criteria at baseline to be qualified as ‘recovered’ after the trial. The post-hoc definition of recovery was based on very loose criteria (mean + 2 SD) yielding impressive ‘recovery rates’. No less than 63% of patients in the FITNET arm met all four recovery criteria: CIS F $<$ 40, CHQ PF \geq 85,

school absence \leq 10%, and “completely recovered” or “much better but still experiencing some symptoms”). However, since the eligibility criteria (CIS F \geq 40, CHQ PF \leq 85 or school absence \geq 15%) border on the ‘recovery’ criteria (CIS F $<$ 40, CHQ PF \geq 85 and school absence \leq 10%), a minimal improvement was sufficient to meet three out of four recovery criteria. Besides the cut-off scores used to define recovery don’t come close to ‘normal values’. For example, a CIS F score $<$ 40 is sufficient to ‘recover from fatigue’, while a CIS F score \geq 35 implies ‘severe fatigue’ in other studies by the same research group [35,59] and a CIS F score \leq 27 is a “level of fatigue comparable to healthy people” [35]. One could argue that these latter cut-off scores relate to adults, not to adolescents. However, another study by the same research group [31] used a cut-off score $<$ 35,7 for CIS F to define ‘clinically significant improvement’ in adolescents. Thus, a patient could ‘recover’ in the FITNET trial [36] without ‘clinically significant improvement’ [31]. A t-test comparison showed that patients ‘recovered’ after FITNET had significantly lower school attendance and worse CHQ PF scores than their healthy peers [39].

When employing more strict criteria (mean + 1 SD) the recovery rate in the FITNET group dropped substantially (from 63% to 36%). However even these ‘strict criteria’ were not very strict. This is illustrated by the observation that a cut-off score of $<$ 35 for CIS F is still very high, especially for adolescent males [60]. The effect of FITNET and ‘usual care’ on the second objective measure mentioned in the protocol [37], physical activity level measures by an actometer, wasn’t reported. Another study by the same research group [29] suggests that a decrease of the school absence could be due to ‘activity substitution’, since an improvement of subjective measures after CBT+ isn’t reflected by an increase of activity levels. In addition, as mentioned before, CBT+ doesn’t yield better cognitive test scores [23], which is relevant, because the IQ of adolescent CFS patients is 8 points lower than their healthy peers, despite a subgroup of patients had already switched to a lower school level. The finding by Nijhoff and others [50] that the IQ wasn’t correlated with school absence in neither the patient group nor in the control group suggests that “IQ is affected in CFS adolescents independently of absence from school” [50] and that the profound fall in IQ isn’t resolved by a reduction in school absence.

Finally, a FITNET follow-up [38] showed the differences in recovery rates (based on the mean + 2 SD criterion) between FITNET and ‘usual care’ vanished over time.

Analysis

Participants

The diagnostic criteria used to select participants in the five studies are CFS [4] in three studies [31-36], CF, erroneously labeled CFS, in one study [7] and a mixed set of CF and CFS [4] in another study [30]. Looking at the type of comorbidities reported, one could question the correct application of the CFS [4] diagnostic criteria in the Knoop *et al.* study [35]. Importantly, no study investigated the effect of CBT+ in a group of patients diagnosed with ME, defined by the original criteria [1-3,11] or the recently proposed International Consensus Criteria [10]. Conflating groups makes it impossible to conclude about a specific group. It affects the validity of a study. Since chronic fatigue (CF) is by far insufficient to meet the diagnosis CFS [4] and the diagnostic criteria for CFS [4] and ME [1-3,11] define distinct and only partially overlapping patient groups [13] (Figure 1), the effects of CBT+ in CF cannot be generalized to CFS, and the effects of CBT+ in CFS cannot be generalized to ME.

Outcomes

Subjective outcomes

Studies into the effects of CBT+ are characterized by a great variety of non-specific subjective measures, e.g. CIS F scores for fatigue, SIP 8 and Karnofsky scores for disability, SF-36 PF and CHQ PF scores for physical functioning. In some cases, the effects of measures to select patients and define caseness were not included in the definitions of improvement or recovery in the same study. For example, clinically significant improvement of the SIP 8 score, one of the two criteria to select patients for the trial, was not reported by Prins *et al.*, 2001 [7]. The recovery rates, 'recovery from CFS' and recovery according to "the most conservative definition of recovery", reported by Knoop *et al.*, 2007 [35] don't include a cut-off criterion for SIP 8, while the SIP 8 score is, besides the CIS F score, one of the two criteria to define CFS. While all studies [7-36] report positive effects at the group level on one or more subjective measures, e.g. fatigue (CIS F) and disability (SIP 8), these effects, if present, are by far insufficient to achieve 'normal values' as defined in other studies by the research group or found in health surveys [55]. In the Bazelmans *et al.*, 2005 trial [30] for example, the modest effect of group CBT+ on the mean CIS F score was largely insufficient to achieve 'normal levels' as defined in other studies of the same research group (≤ 27 [35]). The mean CIS F score after CBT+ would qualify patients as "severely fatigued" by the research group (≥ 40 [7], ≥ 35 [35]). At the individual level, impressive recovery and improvement rates reported by the studies are based on very loose criteria (cut-off thresholds) for subjective measures. For example, the 69% recovery rate reported by Knoop *et al.*, 2007 [35] is based on cut-off scores of < 35 for CIS F (fatigue) and < 700 for SIP 8 (disability), thresholds far above the 'normal levels' as calculated in the same study (CIS F ≤ 27 and SIP 8 ≤ 203). In the Nijhof *et al.*, 2012 trial [36] the eligibility criteria (CIS F ≥ 40 , CHQ PF ≤ 85 or school absence $\geq 15\%$) are on border on the 'recovery' criteria (CIS F < 40 , CHQ PF ≥ 85 and school absence $\leq 10\%$) which implies a minimal improvement is sufficient to be 'recovered'.

Table 3 summarizes the cut-off thresholds for all measures used to define caseness (eligibility criteria) and to define improvement or recovery in one or more studies. The variance of the three measures most often used (CIS F, SIP 8 and SF-36 PF) are illustrated in Figures 2, 3 and 4. As can be seen in Table 3, there is a great variance of cut-off scores used to define caseness, improvement, and recovery. Cut-off criteria for recovery often border on entry criteria, which means a negligible improvement at subjective measures is sufficient to be qualified as 'recovered'. For example, in Nijhof *et al.*, 2002 [36], a patient is diagnosed as a CFS patient with "severe fatigue" and "severe impairment" at the start of the trial if CIS F ≥ 40 and CHQ-CF87 ≤ 85 . The corresponding criteria to be qualified as 'recovered' are CIS F < 40 and CHQ-CF87 ≥ 85 . Importantly, Knoop *et al.*, 2007 [35] qualifies patients with a CIS F score ≥ 35 as "severely fatigued". In some cases, patients are labeled as 'recovered' while not meeting the criteria for clinical improvement in another study. One of the criteria to be qualified as 'recovered' in Nijhoff *et al.*, 2012 [36], is a school attendance $\geq 90\%$, while Stulemeijer *et al.*, 2005 [31] requires 100% school attendance for clinical improvement. Sometimes threshold scores for subjective measures used to define improvement and recovery in studies even show overlap with cut-off scores for CFS caseness in one or more of other studies by the same research group. For example, a score of 61 for the SF-36 PF criterion is sufficient to be qualified as 'recovered' after CBT+ [61], while this score indicates "severe impairment" in other studies [31,62]. The cut-off scores used for subjective measures to

define success (improvement or recovery) fully determine the (positive) outcomes of a study. This accounts for the large variance in 'success rates' reported by the authors. As has been stated by Knoop *et al.*, 2007 [35]: "Recovery is a construction". Subjective measures are sensitive to various psychological effects, illustrated by the finding that 32% of the patients in the non-intervention group reported clinical significant improvement afterwards. For that reason, definitions of improvement recovery should include objective measures.

Objective outcomes

Looking at the objective measures, the proof of clinical improvement after CBT+ is lacking. Improvement on subjective measures, which are sensitive to placebo effects [35], response bias [20], buy-in effects [63], and other psychological effects, isn't reflected by an improvement of objective measures. This is illustrated by the observation that the five studies investigated had no effect on important objective measures, e.g. number of hours worked [7,30] or activity levels [7,31,35]. The lack of effect on the number of hours worked and activity levels is very relevant since "work rehabilitation" is the final goal of CBT+ according to the research group is [7,28]. "Work rehabilitation" implies increased activity levels. School absence is the only measure for which a positive effect was found in the five hallmark studies [31,36]. However, as argued, CBT+ doesn't yield a relevant increase in activity levels [29], which suggests reduced school absence is a consequence of 'activity substitution'. Moreover, as the research group themselves observed, a substantial drop in IQ in adolescent CFS patients is not a consequence of high school absence [50], and cognitive test scores don't improve by CBT+ despite subjective 'improvement' [23].

Discussion

The claim that CBT+ is an evidence-based treatment for CFS isn't supported by the data of five hallmark studies into the effects of CBT+ [7-36]. Although the studies showed (modest) improvement on one or more subjective measures, the improvement wasn't sufficient to achieve 'normal levels' defined by the research group. Various other subjective measures didn't significantly improve when compared to non-intervention, e.g. System Checklist (SCL-90) scores and quality of live (EQL-5D) scores in Prins *et al.*, 2001 [7], and the prevalence of five out of nine CFS symptoms in Stulemeijer *et al.*, 2005 [31]. Sometimes subjective measures even worsened, e.g. SIP 8 (impairment) scores in Bazelmans *et al.*, 2005 [30]. Moreover, except for school absence in two trials [31-36], the effect of CBT+ on objective measures was nihil. A reduction of school absence can plausibly be explained by 'activity substitution' and likely doesn't result in an improved cognitive test scores or an increase of the (low) IQ [50]. The negative effects of CBT+ on the health status, as implicated by various patient surveys [64-71] and a Spanish randomized controlled trial [72], were not investigated in these studies. One study examining the detrimental effects of CBT+ in 212 CFS patients [33], a secondary analysis of Prins *et al.*, 2001 [7], Stulemeijer *et al.*, 2005 [31], and Knoop *et al.*, 2007 [35], observed "no predictors of symptom deterioration specific to [CBT+]." But almost half of the patients in this analysis were suffering from CF [7], and didn't meet the diagnosis CFS [4].

The claim that CBT+ would be effective for ME [1-3] isn't substantiated, since none of the studies (or other trials) investigated the effect of CBT, GET or CBT+ in a group solely consisting of ME [1-3] patients or reported outcomes for CBT+ in the ME patient [1-3] subgroup. If patients are primarily selected by the diagnosis CFS [4], the findings of the CBT+ studies are insufficient to substantiate the claim that CBT+ is effective for ME [1-3], since ME and CFS are distinctive clinical entities [13].

Table 3. Cut-off scores for inclusion in studies, 'normal values', improvement and recovery in the five studies analyzed.

| Measure | Study | Inclusion criteria | 'Normal values' | Criteria for improvement | Criteria for recovery |
|-------------------|-------------------------------|--|--|---|---|
| CHQ-CF87 | Nijhof et al., 2012 [36] | CHQ-CF87 ≤85 or school attendance ≤85% ("functional impairment") | | | CHQ-CF87 ≥85 |
| CIS F | Prins et al., 2001 [7] | CIS F ≥40 ("severe fatigue") | | CIS F ≤36 (reliable change index >1.64, p<0.05) | |
| | Bazelmans et al., 2005 [30] | CIS F ≥35 | | | |
| | Stulemeijer et al., 2005 [31] | CIS F ≥40 ("severe fatigue") | | CIS F <35.7 (reliable change index >1.96) | |
| | Knoop et al., 2007 [35] | CIS F ≥35 ("severely fatigued") | CIS F ≤27 ("fatigue comparable to healthy people") | | CIS F <35 ("Recovery from CFS") |
| | | | | | CIS F ≤27 ("most comprehensive definition of recovery") |
| | Nijhof et al., 2012 [36] | CIS F ≥40 ("severe fatigue") | | | CIS F <40 |
| School attendance | Stulemeijer et al., 2005 [31] | | | School attendance: 100% | |
| | Nijhof et al., 2012 [36] | School attendance ≤85% or CHQ-CF87 ≤85 ("functional impairment") | | | School attendance ≥90% |
| SF36-PF | Stulemeijer et al., 2005 [31] | SF-36 PF ≤65 ("severe functional impairment") | | SF-36 PF ≥75 or change in SF-36 PF ≥50 | |
| | Knoop et al., 2007 [35] | | SF-36 PF ≥80 ("no physical disability") | | |
| SIP 8 | Prins et al., 2001 [7] | SIP 8 ≥800 ("severe functional impairment") | | | |
| | Bazelmans et al., 2005 [30] | SIP 8 ≥700 | | | |
| | Stulemeijer et al., 2005 [31] | | | | |
| | Knoop et al., 2007 [35] | SIP 8 ≥700 ("severely disabled") | SIP 8 ≤203 ("no disabilities in all domains") | | SIP 8 <700 ("Recovery from CFS") |
| | Nijhof et al., 2012 [36] | | | | |

Note: The measures mentioned in Table 3 are described in Table 2.

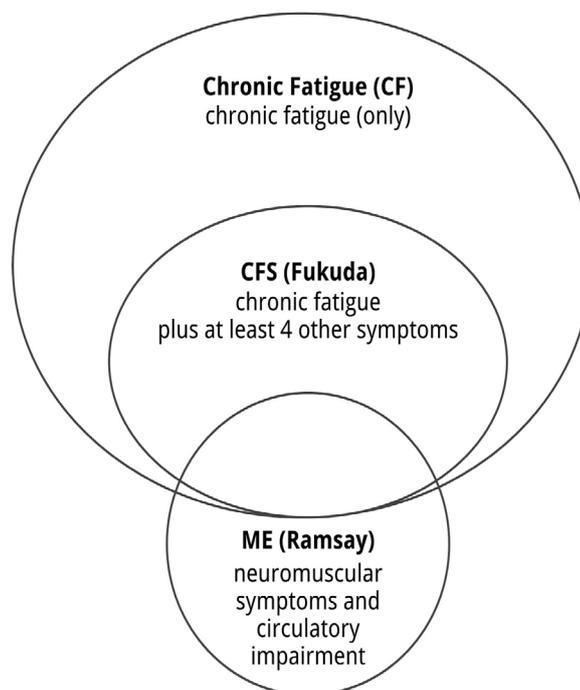


Figure 1. CF, CFS and ME are distinct clinical entities.

Note: Figure 1 only includes case criteria relevant to this study.

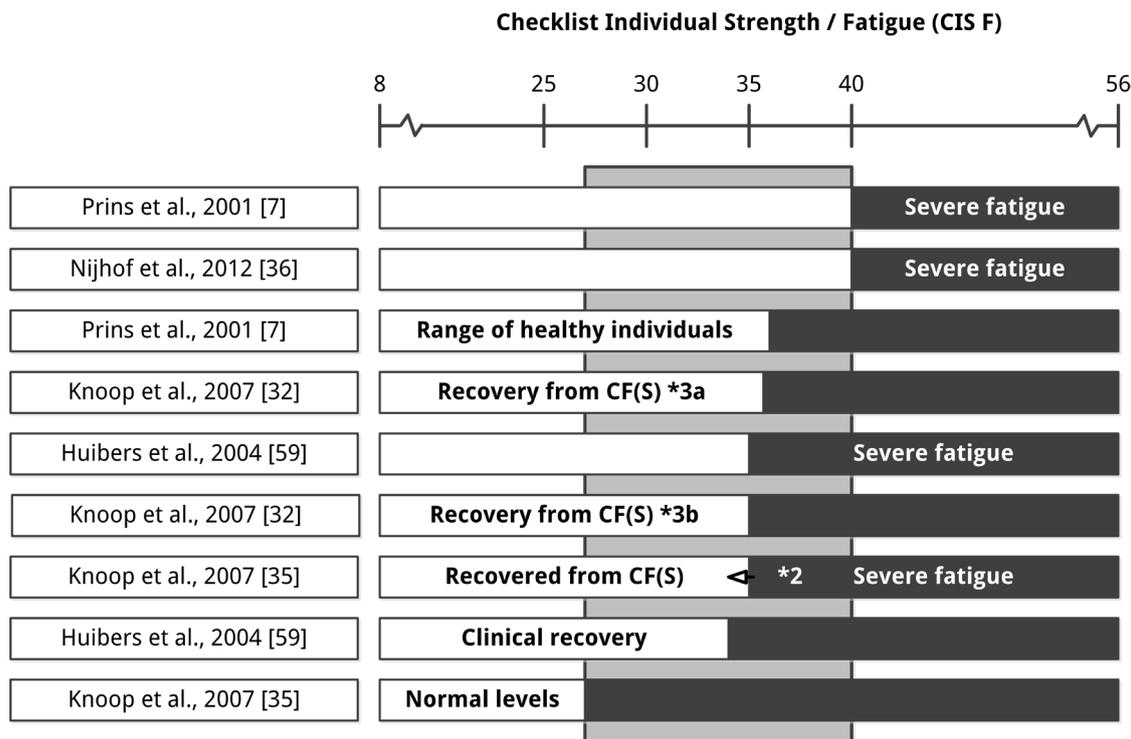


Figure 2. Threshold scores for CIS F (fatigue) to define caseness, improvement and recovery in studies published by the Dutch CF expertise group

*2 a small improvement is sufficient to recover on this aspect

*3 a) adolescents, b) adults

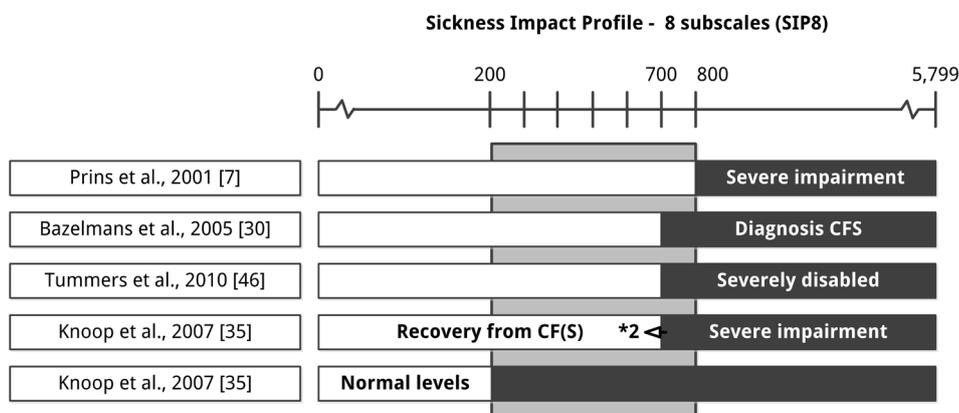


Figure 3. Threshold scores for SIP 8 (impairment) to define caseness, improvement and recovery in studies published by the Dutch CF expertise group.

*2 a small improvement is sufficient to recover on this aspect

The studies suffer from methodological flaws relating to the participants, the intervention, the methods and the outcomes. Issues relating to participants include high rates of self-selection, since “CFS patients are skeptical of psychological interventions” [35], ‘moderate’ and ‘severe cases’ not being able to participate, the use of strongly varying subjective criteria to be eligible to participate, substantial numbers of eligible patients refusing to participate [7], and substantial drop-out rates. Methodological flaws inherently associated with the therapy include the lack of ‘blinding’, and bias and placebo effects, especially when only subjective measures are employed. Shortcomings specifically relating to the methods used in the five studies include the lack of randomisation [30], the lack of a control group [35], the diversity of control conditions, the lack of methods to assess adverse effects, and,

very importantly, not reporting data with regard to objective measures, e.g. Nijhof *et al.*, 2012 [36], or reporting the relevant data years later, e.g. Prins *et al.*, 2001[7]. Methodological shortcomings related to the outcomes of the studies include the use of strongly varying definitions of caseness, improvement and recovery, which means that a patient can ‘recover’ in one study while being qualified a patient (with severe fatigue and impairment) by another study, the use of subjective measures (showing improvement which is not reflected by objective improvement) and defining recovery and (clinical) improvement post-hoc.

In this context it is impossible to address all issues extensively. But considering the number of methodological shortcomings, it is evident that, in addition to the fact that the claims that CBT+ results in clinical

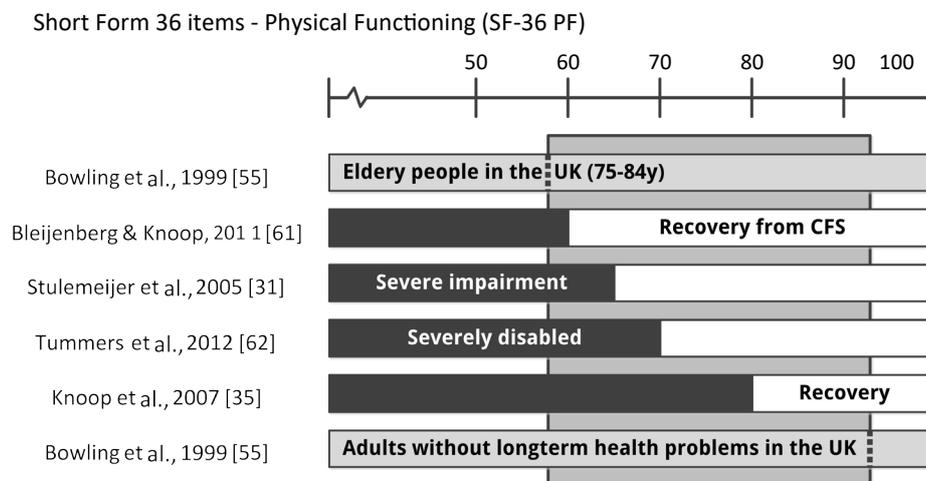


Figure 4. Threshold scores for SF-36 PF (physical functioning) to define caseness, improvement and recovery in studies published by the Dutch CF expertise group and a population study. ■ Reflects the mean score

significant improvement or recovery in ME and CFS and that CBT+ is safe cannot be substantiated by the outcomes, the outcomes are based on studies suffering from serious methodological flaws.

Conclusion

The claim that CBT+ is an effective treatment for CFS isn't supported by the data of five Dutch hallmark studies [7-36]. While some studies show positive effects of CBT+ on subjective measures, e.g. fatigue (CIS F) and physical functioning (SF-36 PF), these effects are insufficient to achieve 'normal values' according to criteria defined by the research group involved and others. CBT+ has no effect on all objective measures, except for one: school absence. The effect on activity levels and number of hours worked is nihil. Looking at the effect of CBT+ on activity levels, the effect on school absence is most likely a consequence of 'activity substitution'. The findings of our analysis are in concordance with the outcomes of the PACE trial [8,73-79]: the (modest) effects of CBT+ on subjective measures are by large insufficient to reach normal levels, CBT+ has no effect on objective measures. All in all, CBT and GET cannot be qualified as curative therapies for CFS [4], let alone ME [1-3,11]. Moreover, there are indications that CBT+ can have detrimental effects.

References

- Ramsay AM (1988) Myalgic Encephalomyelitis and postviral fatigue states: the saga of Royal Free disease.
- Parish JG (1978) Early outbreaks of 'epidemic neuromyasthenia'. *Postgrad Med J* 54: 711-717. [Crossref]
- Dowsett EG, Ramsay AM, McCartney RA (1990) Myalgic Encephalomyelitis - A persistent enteroviral infection? *Postgrad Med J* 66: 526-530. [Crossref]
- Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, et al. (1994) The chronic fatigue syndrome: a comprehensive approach to its definition and study. *Ann Intern Med* 121: 953-959. [Crossref]
- Twisk FNM (2014) The status of and future research into Myalgic Encephalomyelitis and chronic fatigue syndrome: the need of accurate diagnosis, objective assessment, and acknowledging biological and clinical subgroups. *Front Physiol* 5: 109. [Crossref]
- Holgate ST, Komaroff AL, Mangan D, Wessely S (2011) Chronic fatigue syndrome: understanding a complex illness. *Nat Rev Neurosci* 12: 539-544. [Crossref]
- Prins JB, Bleijenberg G, Bazelmans E (2001) Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 357: 841-847. [Crossref]
- White PD, Goldsmith KA, Johnson AL (2011) Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial. *Lancet* 377: 823-836. [Crossref]
- Dowsett EG (1999) Redefinitions of ME/CFS - A 20th Century Phenomenon. ME ForT campaign newsletter.
- Carruthers BM, van de Sande MI, De Meirleir KL, Klimas NG, Broderick G, et al. (2011) Myalgic encephalomyelitis: international consensus criteria. *J Intern Med* 270: 327-338. [Crossref]
- Ramsay AM and Dowsett EG (1992) Myalgic Encephalomyelitis: Then and now. In: BM Hyde, J Goldstein, P Levine (editors), *The Clinical and Scientific Basis of Myalgic Encephalomyelitis*
- Institute of Medicine (2015) *Beyond Myalgic Encephalomyelitis/chronic fatigue syndrome: redefining an illness.* [Crossref]
- Twisk FNM (2016) Replacing Myalgic Encephalomyelitis and chronic fatigue syndrome with Systemic Exercise Intolerance Disease is not the way forward. *Diagnostics (Basel)* 6: 10. [Crossref]
- Sharpe MC, Archard LC, Banatvala JE (1991) Chronic fatigue syndrome: guidelines for research. *J R Soc Med* 84: 118-121. [Crossref]
- Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, et al. (1993) Development of a fatigue scale. *J Psychosom Res* 37: 147-153 [Crossref]
- Vercoulen JHMM, Alberts M and Bleijenberg G (1999) De Checklist Individual Strength (CIS) [Article in Dutch. *Gedragstherapie* 32: 131-136. [Crossref]
- Ware JE Jr, Sherbourne CD (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 30: 473-483. [Crossref]
- Jacobs HM, Luttk A, Touw-Otten FW (1990) The sickness impact profile; results of an evaluation study of the Dutch version De 'Sickness Impact Profile: Resultaten van een valideringsonderzoek van de Nederlandse versie. *Ned Tijdschr Geneesk* 134: 1950-1954.
- Blease C and Kirsch I (2016) The placebo effect and psychotherapy: implications for theory, research, and practice. *Psychol Conscious (Wash D C)* 3: 105-107.
- Furnham A (1986) Response bias, social desirability and dissimulation. *Personal Individ Differ* 7: 385-400
- Munder T, Brüttsch O, Leonhart R, Gerger H, Barth J (2013) Researcher allegiance in psychotherapy outcome research: an overview of reviews. *Clin Psychol Rev* 33: 501-511. [Crossref]
- Weinstein AA, Drinkard BM, Diao G (2009) Exploratory analysis of the relationships between aerobic capacity and self-reported fatigue in patients with rheumatoid arthritis, polymyositis, and chronic fatigue syndrome. *PM R* 1: 620-628. [Crossref]
- Knoop H, Prins JB, Stulemeijer M, et al. (2007) The effect of cognitive behaviour therapy for chronic fatigue syndrome on self-reported cognitive impairments and neuropsychological test performance. *J Neurol Neurosurg* 78: 434-436. [Crossref]
- Twisk FNM (2014) A definition of recovery in myalgic encephalomyelitis and chronic fatigue syndrome should be based upon objective measures. *Qual Life Res* 23: 2417-2418. [Crossref]

25. Gezondheidsraad (2005) Het chronische-vermoeidheidssyndroom. Publication Number 2005/02, Gezondheidsraad, ISBN: 90-5549-554-9, Den Haag (The Netherlands). [[Crossref](#)]
26. Gezondheidsraad (2007) Verzekeringsgeneeskundige protocollen: chronischevermoeidheidssyndroom, lumbosacral radiculair syndroom. Publication Number 2007/12, Den Haag (The Netherlands) [[Crossref](#)]
27. CBO (2013) Richtlijn diagnose, behandeling, begeleiding en beoordeling van patiënten met het chronisch vermoeidheidssyndroom (CVS). [[Crossref](#)]
28. Bleijenberg G, Prins J, Bazelmans E (2003) Cognitive-behavioral therapies. In: Jason LA, Fennel PA, Taylor RR (editors). Handbook of chronic fatigue syndrome, John Wiley, New Jersey.
29. Wiborg JF, Knoop H, Stulemeijer M (2010) How does cognitive behaviour therapy reduce fatigue in patients with chronic fatigue syndrome? The role of physical activity. *Psychol Med* 40: 1281-1287. [[Crossref](#)]
30. Bazelmans E, Prins JB, Lulofs R (2005) Cognitive behaviour group therapy for chronic fatigue syndrome: a non-randomised waiting list controlled study. *Psychother Psychosom* 74: 218-224. [[Crossref](#)]
31. Stulemeijer M, de Jong LWAM, Fiselier TJW (2005) Cognitive behaviour therapy for adolescents with chronic fatigue syndrome: randomised controlled trial. *BMJ* 330: 14. [[Crossref](#)]
32. Knoop H, Stulemeijer M, Prins JB (2007) Is cognitive behaviour therapy for chronic fatigue syndrome also effective for pain symptoms? *Behav Res Ther* 45: 2034-2043. [[Crossref](#)]
33. Heins MJ, Knoop H, Prins JB (2010) Possible detrimental effects of cognitive behaviour therapy for chronic fatigue syndrome. *Psychother Psychosom* 79: 249-256. [[Crossref](#)]
34. Knoop H, Stulemeijer M, de Jong LW (2008) Efficacy of cognitive behavioral therapy for adolescents with chronic fatigue syndrome: long-term follow-up of a randomized, controlled trial. *Pediatrics* 121: e619-e625. [[Crossref](#)]
35. Knoop H, Bleijenberg G, Gielissen MF (2007) Is a full recovery possible after cognitive behavioural therapy for chronic fatigue syndrome? *Psychother Psychosom* 76: 171-176. [[Crossref](#)]
36. Nijhof SL, Bleijenberg G, Uiterwaal CS (2012) Effectiveness of internet-based cognitive behavioural treatment for adolescents with chronic fatigue syndrome (FITNET): a randomised controlled trial. *Lancet* 379: 1412-1418. [[Crossref](#)]
37. Nijhof SL, Bleijenberg G, Uiterwaal CS (2011) Fatigue In Teenagers on the interNET - the FITNET Trial. A randomized clinical trial of web-based cognitive behavioural therapy for adolescents with chronic fatigue syndrome: study protocol. *BMC Neurol* 11: 23. [[Crossref](#)]
38. Nijhof SL, Priesterbach LP, Uiterwaal CS, Bleijenberg G, Kimpen JL, et al. (2013) Internet-based therapy for adolescents with chronic fatigue syndrome: long-term follow-up. *Pediatrics* 131: e1788-1795. [[Crossref](#)]
39. Crawford J (2012) Internet-based CBT for adolescents with chronic fatigue syndrome. *Lancet* 380: 561-562 [[Crossref](#)]
40. Kindlon T (2012) Internet-based CBT for adolescents with chronic fatigue syndrome. *Lancet* 380: 561-562 [[Crossref](#)]
41. Crawley EM (2012) Internet-based cognitive behavioural therapy (FITNET) is an effective treatment for adolescents with chronic fatigue syndrome. *Arch Dis Child Educ Pract Ed* 97: 238. [[Crossref](#)]
42. White PD, Chalder T (2012) Chronic fatigue syndrome: treatment without a cause. *Lancet* 379: 1372-1373. [[Crossref](#)]
43. Karnofsky DA, Abelmann WH, Craver LF (1948) Use of the nitrogen mustards in the palliative treatment of carcinoma - with particular reference to bronchogenic carcinoma. *Cancer* 1: 634-656
44. Schaafsma J and Osoba D (1994) The Karnofsky Performance Status Scale re-examined: a cross-validation with the EORTC-C30. *Qual Life Res* 3: 413-424. [[Crossref](#)]
45. Knoop H, Bleijenberg G and van der Gulden J (2011) Het chronische vermoeidheidssyndroom: aandacht voor werk en werkhervatting *Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde* 2011: 116-119.
46. Tummers M, Knoop H, Bleijenberg G (2010) Effectiveness of stepped care for chronic fatigue syndrome: a randomized noninferiority trial. *J Consult Clin Psychol* 78: 724-731. [[Crossref](#)]
47. Jason LA, Porter N, Hunnell J, Brown A, Rademaker A, et al. (2011) A natural history study of chronic fatigue syndrome. *Rehabil Psychol* 56: 32-42. [[Crossref](#)]
48. Hill NF, Tiersky LA, Scavalla VR, Lavietes M, Natelson BH (1999) Natural history of severe chronic fatigue syndrome. *Arch Phys Med Rehabil* 80: 1090-1094. [[Crossref](#)]
49. Vercoulen JHMM, Swanink CM, Fennis JF (1996) Prognosis in chronic fatigue syndrome: a prospective study on the natural course. *J Neurol Neurosurg Psychiatry* 60: 489-494. [[Crossref](#)]
50. Nijhof LN, Nijhof SL, Bleijenberg G, Stellato RK, Kimpen JL, et al. (2016) The impact of chronic fatigue syndrome on cognitive functioning in adolescents. *Eur J Pediatr* 175: 245-252. [[Crossref](#)]
51. Cockshell SJ, Mathias JL (2010) Cognitive functioning in chronic fatigue syndrome: a meta-analysis. *Psychol Med* 40: 1253-1267. [[Crossref](#)]
52. Torrance N, Smith BH, Lee AJ (2009) Analysing the SF-36 in population-based research. A comparison of methods of statistical approaches using chronic pain as an example. *J Eval Clin Pract* 15: 328-334. [[Crossref](#)]
53. Rose MS, Koshman ML, Spreng S (1999) Statistical issues encountered in the comparison of health-related quality of life in diseased patients to published general population norms: problems and solutions. *J Clin Epidemiol* 52: 405-412. [[Crossref](#)]
54. Aaronson NK, Muller M, Cohen PD (1998) Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiol* 51: 1055-1068. [[Crossref](#)]
55. Bowling A, Bond M, Jenkinson C (1999) Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *J Public Health Med* 21: 255-270. [[Crossref](#)]
56. Castro-Marrero J, Faro M, Aliste L (2017) Comorbidity in chronic fatigue syndrome/Myalgic Encephalomyelitis: a nationwide population-based cohort study. *Psychosomatics* 58: 533-543. [[Crossref](#)]
57. Landgraf JM and Abetz L (1997) Functional status and well-being of children representing three cultural groups: initial self-reports using the CHQ-CF87. *Psychol Health* 12: 839-854.
58. Raat H, Landgraf JM, Bonsel GJ (2002) Reliability and validity of the child health questionnaire-child form (CHQ-CF87) in a Dutch adolescent population. *Qual Life Res* 11(6): 575-581. [[Crossref](#)]
59. Huibers MJ, Beurskens AJ, van Schayck CP (2004) Efficacy of cognitive-behavioural therapy by general practitioners for unexplained fatigue among employees: Randomised controlled trial. *Br J Psychiatry* 184: 240-246. [[Crossref](#)]
60. ter Wolbeek M, van Doornen LJ, Kavelaars A, Heijnen CJ (2006) Severe fatigue in adolescents: a common phenomenon? *Pediatrics* 117: e1078-1086. [[Crossref](#)]
61. Bleijenberg G, Knoop H (2011) Chronic fatigue syndrome: where to PACE from here? *Lancet* 377: 786-788. [[Crossref](#)]
62. Tummers M, Knoop H, van Dam, et al. (2012) Implementing a minimal intervention for chronic fatigue syndrome in a mental health centre: a randomized controlled trial. *Psychol Med* 42(10): 2205-2215. [[Crossref](#)]
63. Vos-Vromans DC, Huijnen IP, Rijnders LJ (2016) Treatment expectations influence the outcome of multidisciplinary rehabilitation treatment in patients with CFS. *J Psychosom Res* 83: 40-45. [[Crossref](#)]
64. ME Association (2010) Managing my ME. What people with ME/CFS and their carers want from the UK's health and social services. Gawcott
65. ME Association (2015) ME/CFS Illness Management Survey Results: No decisions about me without me.
66. Action for ME (2001) Severely neglected, London.
67. Action for ME (2007) Scotland M.E./CFS scoping exercise report.
68. Bjørkum T, Wang CE and Waterloo K (2009) Patients' experience with treatment of chronic fatigue syndrome [Article in Norwegian] Pasienterfaringer med ulike tiltak ved kronisk utmattelsessyndrom. *Tidsskr Nor Laegeforen* 129: 1214-1216. [[Crossref](#)]
69. Bringsli GJ, Gilje A and Getz Wold BK (2014) The Norwegian ME Association National Survey
70. Koolhaas MP, de Boorder H and van Hoof E (2008) Cognitieve gedragstherapie bij het chronische vermoeidheidssyndroom (ME/CVS) vanuit het perspectief van de patiënt [Article in Dutch]. *Medisch Contact*.
71. De Kimpe A, Crijnen B, Kuijper J, et al. (2016) [Zorg voor ME - Enquête onder ME-patiënten naar hun ervaringen met de zorg in Nederland 2016] [Article in Dutch]

72. Núñez M, Fernández-Solà J, Nuñez E (2011) Health-related quality of life in patients with chronic fatigue syndrome: group cognitive behavioural therapy and graded exercise versus usual treatment. A randomised controlled trial with 1 year of follow-up. *Clin Rheumatol* 30: 381-389. [[Crossref](#)]
73. McCrone P, Sharpe M, Chalder T (2012) Adaptive pacing, cognitive behaviour therapy, graded exercise, and specialist medical care for chronic fatigue syndrome: a cost-effectiveness analysis. *PLoS One* 7: e40808. [[Crossref](#)]
74. White PD, Goldsmith KA, Johnson AL (2013) Recovery from chronic fatigue syndrome after treatments given in the PACE trial. *Psychol Med* 43: 2227-2235. [[Crossref](#)]
75. Sharpe MC, Goldsmith KA, Johnson AL (2015) Rehabilitative treatments for chronic fatigue syndrome: long-term follow-up from the PACE trial. *Lancet Psychiatry* 2: 1067-1074. [[Crossref](#)]
76. Chalder T, Goldsmith KA, White PD (2015), Rehabilitative therapies for chronic fatigue syndrome: a secondary mediation analysis of the PACE trial. *Lancet Psychiatry* 2: 141-152. [[Crossref](#)]
77. Wilshire C, Kindlon T, Matthees A (2017) Can patients with chronic fatigue syndrome really recover after graded exercise or cognitive behavioural therapy? A critical commentary and preliminary re-analysis of the PACE trial. *Fatigue: Biomedicine, Health & Behavior* 5: 43-56. [[Crossref](#)]
78. Geraghty KJ (2017) 'PACE-Gate': When clinical trial evidence meets open data access. *J Health Psychol* 22: 1106-1112. [[Crossref](#)]
79. Twisk F (2016) PACE: CBT and GET are not rehabilitative therapies. *Lancet Psychiatry* 3: e6. [[Crossref](#)]