Is cognitive behaviour therapy with a graded activity protocol effective for chronic fatigue syndrome and Myalgic Encephalomyelitis? An analysis of five Dutch hallmark studies.

Reference


Introduction

Advocates of the (bio)psychosocial explanatory model for 'ME/CFS' and the 'behavioral interventions' based on this model (CBT: cognitive behavioral therapy and GET: gradual exercise therapy) often refer to the results of the PACE trial and studies from the Netherlands.

The prestigious PACE trial has been controversial from the very first moment. Recently, the Journal of Health Psychology published a special issue to address the methodical shortcomings and disappointing outcomes of the PACE trial.

Various analyzes of the PACE trial and follow-up studies have made it clear
• that the PACE trial suffers from various serious methodological flaws,
• that the patient group studied were not CFS patients, let alone people with ME, but people with chronic fatigue (CF),
• that the effects of CBT and GET on subjective measures were (very) modest,
• that, looking at the PACE trial protocol, there is no question of recovery, and
• there was no objective change in the (hopeless) situation of patients.

A second important pillar of the 'evidence-base' for CBT and CBT+ in 'ME/CFS' are the studies of the Dutch Expertise Center for Chronic Fatigue (NKCV). CBT+ stands for CBT combined with/aimed at an incremental increase of activities.
Is the claim of the NKCV that CBT+ is effective for CFS and ME justified by the data reported by five high-profile studies of the NKCV?

For a study recently published in General Medicine Open, five studies of the NKCV were selected on the basis of four criteria. The studies • are frequently cited in the scientific literature and/or • have, partly due to press releases of the NKCV, received a lot of media attention, • examined CBT+ in different variants (individually or in group sessions, face-to-face or through the internet/e-mail contact, for adults or adolescents) and • have a strong influence on medical policies for ME and CFS in the Netherlands.

To establish if the suggested effectiveness of CBT+ for CFS is supported by the (reported) data, the following five studies (and related secondary outcome studies, follow-up studies, commentaries and editorials) were analysed thoroughly:
• the prestigious Prins et al., 2001 trial published in the Lancet,
• the Bazelmans et al., 2005 trial that investigated CBT+ ‘delivered’ in groups,
• the Stulemeijer et al., 2005 trial which studied CBT+ in adolescents,
• the Knoop et al., 2007 study that claimed impressive recovery rates and
• the Nijhof et al., 2012 (FITNET)-trail that studied internet CBT+ in adolescents.

The most important topics in the analysis of the studies were:
• the patient group investigated: ME, CFS and/or chronic fatigue (CF)?
• the subjective and objective measures used to assess the effects of CBT+
• the cut-off threshold scores for the subjective measures, such as ‘fatigue’ and ‘impairment’ used as criteria to define ‘clinical improvement’ or ‘recovery’,
• the consistency of definitions of 'patient', 'improvement' and 'recovery”, and
• the objective effects of CBT+, e.g. effect on the number of hours worked.

The results of that analysis are (briefly) described in the following pages. For a detailed analysis of the studies, you are referred to the article.
The five studies suffer from serious methodical flaws

The five studies examined have various serious methodological shortcomings, relating to the participants, the research methods used and the outcomes.

Methodological shortcomings with regard to participants include:
• self-selection of the participants (patients who do not sign up for these trials), because patients are very skeptical about psychological interventions,
• the observation that 'average' and 'serious cases' were not able to participate,
• the fact that a number of studies do not relate to CFS, but to CF or CFS/CF,
• the observation that CBT+ has never been studied in a group of ME patients,
• the use of strongly fluctuating criteria to define and to select participants,
• the fact that substantial group of eligible patients didn't want to participate, and
• the high percentage of drop-outs/withdrawals during the trials.

Criticism relating to the research methods used include:
• the fact that the NKCV consistently use the label cognitive behavioral therapy (CBT), while gradual expansion of activities is the goal of the therapy (CBT+),
• the fact that patients know that they are subjected to therapy ('lack of blinding'),
• the risk that the outcomes will be (strongly) influenced by, among other things, by a response bias, the placebo effect and the bias of the researchers, especially if (almost) only subjective outcome measures are used,
• the lack of a control group, in particular a non-intervention group, since non-intervention also seem to have a positive effect in many patients, and
• not reporting data or reporting data much later, especially objective outcomes.

Methodological criticism with regard to the outcomes include:
• the use of subjective measures to determine the effect of CBT+, especially since the NKCV stipulates that increased activity and resuming work are the goals,
• the use of incorrect statistical methods to determine 'normal values',
• defining recovery/improvement afterwards ('post hoc') and
• the use of various different definitions of recovery and improvement, so that someone meets the criteria for recovery in one study of the NKCV, while he/she meets the admission criteria for 'severe fatigue' of another study.

In this context it is impossible to review all methodological flaws extensively. In this summary only a number of relevant shortcomings will be discussed.
Some trails studied CFS patients are investigated, other studies investigated CF patients or a mix of CF and CFS patients, but the effect of CBT+ in (a group of) ME patients has never been established.

Two or three (?) trials investigated the effect of CBT+ in a group of CFS patients, a fourth study explored CBT+ in patients with chronic fatigue, while the fifth study evaluated CBT+ in a group of patients with the diagnosis CFS or chronic fatigue.

With regard to one of the three studies into the effects of CBT+ in CFS, serious question marks can be raised about the correctness of the diagnosis of CFS. Some participants experienced ('comorbid') medical/psychological disorders which could plausibly explain 'fatigue', an exclusion criteria for a diagnosis CFS. On the other hand, only a few patients reported other ('comorbid') disorders, such as bowel problems, while 80% of the CFS patients experience co-morbidities.

The effect of CBT+ on ME patients has never been examined or reported (so far). So the effectiveness claim for CBT+ in ME lacks any scientific justification. By definition, ME, a neuromuscular ('polio-like') disease, is not a synonym of CFS.

In the case of ‘ME/CFS’, CBT is not the equivalent to cognitive behavioral therapy, but behavioral therapy combined with gradual expansion of activities, hence: CBT+.

Cognitive behavioral therapy (CBT) is a psychotherapy aimed at 'dysfunctional cognitions' (e.g. 'fear of movement') and subsequent behavior (e.g. avoiding activities).

But in the case of ‘ME/CFS' 'CBT' refers to a combination of cognitive behavioral therapy (CBT) and 'rehabilitation therapy': gradual expansion of the activity levels, which, according to the NKCV, should ultimately result in work resumption.

Looking at recent studies into energy production and metabolism, in particular
• a study showing that the energy production under pressure is largely inadequate,
• a study that found a deviant response of isolated muscle cells to exercise, and
• a study that observed metabolic abnormalities (a survival mechanism?), one could raise several questions regarding the medical ethics of the researchers.

The five studies are largely based on subjective measures which are very sensitive to placebo effect, response bias, bias of researchers etc.

The NKCV employs a wide range of subjective measures in its studies. Subjective measures are incomparable in time and in-between patients.

Moreover, these subjective measures introduce a high risk of a response bias (giving a socially desirable answer), a placebo effect, a buy-in effect (related to the expectations that the patient has before the therapy), the effect of the individual therapist, the therapist-patient relationship, but above all, the effect of bias of the researchers ('researcher allegiance').

For example, in one study the SIP 8 score was one of the two criteria to be eligible. The effect of CBT+ in this study was largely insufficient to achieve 'normal values'. The SIP8 score is not included in the 'general' and 'strict' definition of 'recovery'. These definitions 'justify' the 'recovery rates' that are mentioned in the media (70%).
Studies use widely varying 'cut-off values' for recovery and improvement

The five studies are characterized by a wide range of subjective measures with strongly fluctuating cut-off values for 'clinical improvement' and 'recovery'. The (arbitrary) cut-off values fully determine the outcomes of a study, e.g. recovery rates. As a study by the NKCV stipulates: 'Recovery is a construction'.

The use of highly fluctuating 'cut-off thresholds' for the three most commonly used measures, the CIS F score for 'the level of fatigue' and the SIP 8 and SF-36 PF score for the 'functional/physical limitations', are illustrated in Figure 2, 3 and 4.

The use of various 'cut-off values' in the studies implicate that someone meets one of the criteria for 'recovery' in one study, while he/she meets the eligibility criterion for 'severe fatigue' of another study.

Objective criteria are used very sporadically

While the NKCV poses in one study that "It is important that [patients] jointly (re) formulate a definition of the objective of the treatment." and patients have repeatedly reported that CBT+ has no or a negative effect, the effect of CBT+ on objective measures, such as the level of activity, the number of hours worked and school performance, is rarely determined objectively.

This is very remarkable, to say the least, especially because these objective measures reflect the goals of CBT+ according to the NKCV.

The objective assessment of the effects of CBT+ is also important, because one study conducted by the NKCV observed that self-reported activity levels don't correspond well with the objective determination thereof.

Objective outcomes of the study protocol were not reported or reported much later

Contrary to what is indicated in the study or the protocol of the trial, the effects of CBT+ on a highly relevant objective measure, the activity level, sometimes are not reported at all or are only reported many years later.

Given the lack of impact of CBT+ on activity levels observed in three studies, and the fact that an increased activity level is the goal of CBT+ according to the NKCV, not reporting these data or reporting data nine years later is a scientific 'mortal sin'.

In summary, we can conclude that the five studies of the NKCV are characterized by serious scientific flaws.
Figure 1.
ME and CFS are two distinct, partially overlapping, diagnostic entities

Chronic Fatigue (CF)
chronic fatigue (only)

CFS (Fukuda)
chronic fatigue
plus at least 4 other symptoms

ME (Ramsay)
neuromuscular symptoms and circulatory impairment

Figure 2.
Threshold scores used for the definition of recovery, improvement and caseness (patient selection) in studies conducted by the NKCV (ECCF): CIS F fatigue score

Checklist Individual Strength / Fatigue (CIS F)

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>56</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prins et al., 2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nijhof et al., 2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prins et al., 2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knoop et al., 2007a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huibers et al., 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knoop et al., 2007c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knoop et al., 2007b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huibers et al., 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knoop et al., 2007b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Severe fatigue
- Range of healthy individuals
- Recovery from CF(S) *3a
- Severe fatigue
- Recovery from CF(S) *3b
- Recovered from CF(S) *2
- Severe fatigue
- Clinical recovery
- Normal levels

*2 a small improvement is sufficient to recover on this aspect
*3 a) adolescents, b) adults
Figure 3. Threshold scores used for the definition of recovery, improvement and caseness (patient selection) in studies conducted by the NKCV (ECCF): SIP 8 disability score

![Sickness Impact Profile - 8 subscales (SIP8)](image)

- Prins et al., 2001: Severe impairment
- Bazelmans et al., 2005: Diagnosis CFS
- Tummers et al., 2010: Severely disabled
- Wiborg et al., 2015: Substantial impairment
- Knoop et al., 2007b: Recovery from CF(S) *2* Severe impairment
- Knoop et al., 2007b: Normal levels

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

Figure 4. Threshold scores used for the definition of recovery improvement and caseness (patient selection) in studies by the NKCV (ECCF): SF-36 PF physical functioning score

![Short Form 36 items – Physical Functioning (SF-36 PF)](image)

- Bowling et al., 1999: Elderly people in the UK (75-84y)
- Bleijenberg/Knoop, 2011: Recovery from CFS
- Stulemeijer et al., 2005: Severe impairment
- Tummers et al., 2012: Severely disabled
- Knoop et al., 2012: Recovery
- Bowling et al., 1999: Adults without longterm health problems in the UK

... reflects the mean score
The five studies show that CBT+ is not effective for CFS, the effect of CBT+ for ME was never investigated/reported

Even if we would accept the numerous methodical shortcomings, the claim that CBT+ is effective for CFS and ME cannot be substantiated.

CBT+ has a (modest) positive, no or a negative effect on subjective measures at group level, such as 'fatigue' and 'functional limitations'

CBT+ has a (modest) positive effect on a number of subjective measures, CBT+ has no significant effect on a various other subjective measures, and CBT+ has a negative effect on the primary subjective criterion in one study.

CBT+ has no effect whatsoever on objective measures relevant to the NKCV

With one exception, CBT+ had no effect on the objective measures used, especially activity levels, cognitive test scores, and number of hours worked. The latter is very relevant since work resumption is the goal according to the NKCV.

The only objective positive effect was an improvement of school attendance, that was reported in a study from 2005 and a study from 2012.

However, most likely this positive effect is the result of 'activity substitution', doing a activity, school work, at the expense of another, for example sport, since an NKCV study from 2010 (analysing data of three trials) showed that CBT+ has no substantial effect on the activity levels as assessed by actometers.

In this context it is relevant that other studies of the NKCV demonstrated that CBT+ has no (significant) effect on cognitive test scores that the substantial drop of the IQ is not related to school absenteeism.

Impressive recovery rates are based on very loose, widely varying subjective criteria

The impressive percentages of patients who are supposed to improve or recover (Knoop and colleagues report a recovery rate of 69% in 2007) are based on extremely loose and very strongly fluctuating subjective criteria, for example cut-off thresholds for the CIS F fatigue score and SIP 8 disability score, so that a minimal improvement or a zero effect is sufficient to 'recover', so that improvement is completely inadequate to achieve 'normal values' ('normal values' as defined by the NKCV itself) and so that a 'recovered patient' meets the eligibility criteria of other NKCV studies.

As indicated by Knoop and colleagues, 'recovery is a construction' based on strongly fluctuating 'normal values' for ill-defined notions, like 'fatigue'. Extremely loose 'normal values' for subjective definitions of 'recovery' enable researchers to report very impressive 'recovery percentages'.

In summary, we can conclude that CBT+ (in some cases) has a (small) positive effect on subjective measures, but that this effect is insufficient for patients to be qualified as 'recovered', that CBT+ has no effect on relevant objective indicators, and that the effect of CBT+ in ME was never investigated.