Critical Review

HEALTH PSYCHOLOGY OPEN

Health Psychology Open
January-June 2019: 1–23

DOI: 10.1177/2055102919840614
journals.sagepub.com/home/hpo
SAGE

Cognitive behavioural therapy for myalgic encephalomyelitis/chronic fatigue syndrome is not effective. Re-analysis of a Cochrane review

Mark Vink[1] and Alexandra Vink-Niese[2]

Abstract

Analysis of the 2008 Cochrane review of cognitive behavioural therapy for chronic fatigue syndrome shows that seven patients with mild chronic fatigue syndrome need to be treated for one to report a small, short-lived subjective improvement of fatigue. This is not matched by an objective improvement of physical fitness or employment and illness benefit status. Most studies in the Cochrane review failed to report on safety or adverse reactions. Patient evidence suggests adverse outcomes in 20 per cent of cases. If a trial of a drug or surgical procedure uncovered a similar high rate, it would be unlikely to be accepted as safe. It is time to downgrade cognitive behavioural therapy to an adjunct support-level therapy, rather than a treatment for chronic fatigue syndrome.

Keywords

chronic fatigue syndrome, Cochrane review, cognitive behavioural therapy, myalgic encephalomyelitis

Introduction

For years, the recommended treatments for chronic fatigue syndrome (CFS) have been cognitive behavioural therapy (CBT) and graded exercise therapy (GET). These recommendations have been based on Cochrane reviews (Larun et al., 2017; Price et al., 2008) and a large randomised controlled trial by White et al. (2011), informally referred to as the PACE trial ('*P*acing, graded *A*ctivity, and *C*ognitive behaviour therapy; a randomised *E*valuation'). This trial concluded that CBT and GET were moderately effective treatments, leading to recovery in 22 per cent of patients. Due to its size ($n=640$) and promotion, it has been very influential in the promotion of CBT and GET as effective treatments for CFS (Wilshire et al., 2018b). Recently, a number of re-analyses of the PACE trial, including a special issue of the *Journal of Health Psychology* (Marks, 2017), have raised significant concerns with the published outcomes of the trial. If the PACE trial had not made a significant number of outcome changes, which led to an overlap in entry and recovery criteria, then there would not have been a difference in recovery rate between CBT and GET and the two control groups (no treatment (specialist medical care) and adaptive pacing therapy) (Geraghty, 2017a; Vink, 2016; Wilshire et al., 2018b). Essentially, the recovery rate would have been the same as the natural occurring one (Cairns and Hotopf, 2005). The absence of objective improvement in the PACE trial (fitness and 6-minute walk test (6MWT)) and the increase in illness and unemployment benefits, matched the findings from the evaluation of the use of CBT and GET in the Belgium CFS knowledge centres (Stordeur et al., 2008). As noted by O'Leary (2018), 'although PACE [has] dictated management of ME/CFS across the globe for many years, the study fails to meet basic standards of scientific methodology'. 'Indeed, it is difficult to imagine how such a large-scale investigation could have developed, proceeded and passed through the review process unless its scientific failings

[1]Family and Insurance Physician, Amsterdam, The Netherlands
[2]Independent Researcher, Germany

Corresponding author:
Mark Vink, Family and Insurance Physician, 1096 HZ Amsterdam, The Netherlands.
Email: markvink.md@outlook.com

were actually characteristic of its field'. Analysis of the Dutch FITNET trial of Internet CBT for adolescents (Ghatineh and Vink, 2017), of the Dutch FatiGo trial of multidisciplinary rehabilitation treatment (Vink and Vink-Niese, 2018a) and of five Dutch hallmark CBT studies (Twisk and Corsius, 2017) supported this observation. A recent re-analysis of the Cochrane exercise review for CFS (Vink and Vink-Niese, 2018b) revealed a number of methodological concerns with many of the studies reviewed as part of the Cochrane review of GET for CFS and a lack of objective evidence for improvement in physical function. It also showed that the problems noted by O'Leary are not confined to Dutch studies. O'Leary also concluded that 'the PACE controversy suggests a need to evaluate the scientific credibility of psychosomatic medicine generally'. As such, we carried out an analysis of the Cochrane review of CBT treatment for CFS by Price et al. (2008), to ascertain if this review contained any of the problems identified in Vink and Vink-Niese (2018b), by O'Leary (2018) or Geraghty (2017a) and also to assess whether or not the conclusions of this Cochrane review – that CBT is somewhat effective with moderate size effects – is justified by the data contained within the primary studies included in the review. In our analysis, we concentrated on the objective outcome measures to establish if improvements in self-report (fatigue) translate to observable improvement in objective tests (physical ability, fitness, etc.) as there is an inverse relationship between fatigue and physical activity (Rongen-van Dartel et al., 2014).

The Cochrane CBT review for CFS was set up to determine the effectiveness and acceptability of CBT for patients with CFS alone and in combination with other interventions, compared with any other intervention or control. The review concluded that CBT is effective in reducing fatigue (Price et al., 2008). It included seven randomised controlled studies (RCTs) of CBT for CFS and 820 participants: Lloyd et al. (1993) ($n=90$), Sharpe et al. (1996) ($n=60$), Deale et al. (1997) ($n=60$), Prins et al. (2001) ($n=278$), Whitehead and Campion (2002) ($n=65$), O'Dowd et al. (2006) ($n=153$) and Jason et al. (2007) ($n=114$).

It also included four unpublished studies (Barrett, 1992, $n=$unknown; Stevens et al., 1999b, $n=$unknown; Russell et al., 2001, $n=$unknown; and Strang, 2002, $n=51$) that had not gone through peer review; as well as one study of mindfulness (Surawy et al., 2005), which consisted of three little studies ($n=41$: $18+12+11$) and three fatigue studies which included a minority of CFS patients (28%, 29% and 44%) (King, 1999, published as Ridsdale et al., 2001, $n=45$ of 160; Ridsdale et al., 2004, $n=36$ of 123 and Huibers et al., 2004, $n=66$ of 151). The total number of CFS patients in the review is 1008 ($820+188$; the four unpublished studies not included).

According to Price et al. (2008), there were six trials still in progress (including White et al., 2011), which will help to strengthen the evidence base on CBT interventions for CFS. They wrote that in 2008. These trials should have been published by now and we will analyse them too (Appendix 1). Two studies, Knoop et al. (2007a) and Flo and Chalder (2014), were used by White et al. (2011) as support for their recovery claims. We will therefore analyse these two studies too (Appendix 2). Price et al. (2008) state that CFS has had many names in recent decades, including post-viral fatigue syndrome and myalgic encephalomyelitis (ME), but in the rest of their document, the term CFS is used. The same has been done here to avoid any confusion.

In our analysis, we found 12 areas of concern. We also identify a number of problems in the original studies, including a failure to report harms. Our analysis shows that CBT only leads to a small, short-lived subjective improvement of fatigue in 14 per cent of patients which is not matched by objective improvement. When the objective outcomes of the trials are considered, it is possible to state that CBT is not an effective treatment for CFS.

Twelve areas of concern

1. Reviewer/researcher biases and treatment allegiances

First, selection of the editorial group from the Cochrane Depression, Anxiety and Neurosis Group suggests a lack of neutrality and bias to a particular view of CFS as ME has been classified as a neurological disease by the World Health Organization since 1969, with CFS as an equivalent (WHO, 2010). The four authors of the review are all proponents of the biopsychosocial model, which views CBT as an effective treatment for CFS.

Second, the review itself and 10 of the 11 studies (Jason et al., 2007, the exception; the four unpublished studies not included) were conducted by researchers with an allegiance to a particular model of CFS and to two interventions, CBT and GET, who wanted to prove their own theories. In studies examining more than one treatment approach, the treatment favoured by the researchers tends to outperform other treatments (Luborsky et al., 1999, 2002; Munder et al., 2012). Several factors may contribute to this effect, but one is likely to be the manner in which the non-favoured, comparison treatment is conceptualised and implemented. Often when a treatment is used as a comparison condition, it is implemented in a weaker form then when it is used clinically. Usually investigators do not believe in the effectiveness of the control condition. Consequently, treatments might not be presented to participants as equally likely to lead to improvement. This is especially important when the primary outcomes are self-report measures which can be strongly influenced by patients' expectations. Finally, a researcher's enthusiasm for a particular treatment can also lead them to overinterpret their findings or overlook limitations (Wilshire, 2017).

A review of clinical trials by Lundh et al. (2012) came to the conclusion that industry sponsored drug and device

studies are more often favourable to the sponsor's products than non-industry sponsored drug and device studies due to biases that cannot be explained by standard risk of bias assessment tools. Munder et al. (2013) concluded that a pre-existing belief in or adherence to a particular therapy or treatment may play a similar role in leading to more favourable results in clinical trials.

Before they conducted their research, Chalder and Wessely (co-authors in Deale et al., 1997), Sharpe, O'Dowd, Prins, Surawy, Ridsdale and Bleijenberg (co-author in Prins et al., 2001 and Huibers et al., 2004) are all known to have favoured the approach to the illness being tested. These investigators are deeply committed to the 'unhelpful cognitions' theory of ME/CFS, which they and other colleagues had originated and/or actively promoted. If their trials had failed to show significant improvement and recovery through CBT, that would have undermined the very theories of reversibility to which the investigators have dedicated their careers. Consequently, the risk of latent bias was palpable from the outset (Lubet, 2017). It is notable that the study conducted by a researcher without an allegiance to the model, concluded that none of the four treatment strategies was superior to another treatment strategy in all areas (Jason et al., 2007).

2. A study was excluded that contradicted the main findings

Friedberg and Krupp (1994) was excluded from the Cochrane review because it was a non-randomised controlled trial (Price et al., 2008). All selected participants were offered CBT; those refusing it were assigned to the no treatment group. Consequently, only motivated participants received CBT. This study found that CBT helped to reduce the symptoms of depression, stress and fatigue severity in depressed but not in non-depressed CFS patients.

3. Criteria used in the trials were too broad

As inclusion criteria, four studies (Deale et al., 1997; Prins et al., 2001; Sharpe et al., 1996; Surawy et al., 2005) used the Oxford criteria. Six studies (Huibers et al., 2004; Jason et al., 2007; O'Dowd et al., 2006; Ridsdale et al., 2004; Ridsdale et al., 2001; Whitehead and Campion, 2002) used the Centers for Disease Control and Prevention (CDC) 1994 criteria, also known as the Fukuda Criteria (Fukuda et al., 1994).

The only requirement for the Oxford criteria (Sharpe et al., 1991), is 6 months or more of unexplained fatigue. It was created as an alternative, less strict, operational definition which is essentially chronic fatigue in the absence of neurological signs with psychiatric symptoms as common associated features (David, 1991). The Oxford criteria are untenable because they inappropriately select healthy subjects with mild fatigue and chronic idiopathic fatigue and mislabel them as CFS (Baraniuk, 2017). The American National Institute of Health (NIH) concluded in 2014 that the Oxford criteria are flawed and include people with other conditions, confounding the ability to interpret the science (Green et al., 2014a). Continuing to use the Oxford definition may impair progress and cause harm (Green et al., 2014a; Green et al., 2014b). The Agency for Healthcare Research and Quality (AHRQ) stated that using the Oxford case definition results in a high risk of including patients who may have an alternate fatiguing illness or whose illness resolves spontaneously with time (Smith et al., 2016). Both the NIH and AHRQ recommend that the Oxford definition should be retired.

The Fukuda criteria are the most commonly used criteria for CFS. Patients need to have 6 months or more of unexplained chronic fatigue and fulfil a minimum of four out of eight criteria. However, ME is a multisystem disease including nervous, cardiovascular, endocrine and other involvement, distinguished by severe and prolonged muscle fatigue following trivial exertion. Other characteristics include high morbidity, low mortality, a prolonged relapsing course and variation in symptoms within and between episodes, tending to chronicity (Dowsett et al., 1990). The prestigious American Institute of Medicine (IOM, now the National Academy of Medicine) defined PEM (postexertional malaise) as an exacerbation of some or all of an individual's ME/CFS symptoms that occurs after physical or cognitive exertion and leads to a reduction in functional ability. The IOM concluded that ME/CFS is a systemic exercise intolerance disease (IOM, 2015). This core symptom of CFS is only optional and not compulsory for diagnosis in Fukuda, as it is one of the eight additional criteria (Fukuda et al., 1994); 15 per cent of people labelled by these criteria as having CFS, were in fact healthy people (Friedberg et al., 2000).

The use of the Oxford and the Fukuda criteria in the studies (Lloyd et al., 1993, the exception) means that they may have included patients who did not have CFS, but who were susceptible to the interventions or whose illness resolves spontaneously with time.

4. Problems with the comparison or control group

One study (Surawy et al., 2005) used a waitlist control group for one of its included studies (its other two had no control group) and six studies (Huibers et al., 2004; O'Dowd et al., 2006; Prins et al., 2001; Ridsdale et al., 2004; Sharpe et al., 1996; Whitehead and Campion, 2002) had a no treatment control group, labelled as normal care, specialist medical care, natural course and so on, in which patients could see their own general practitioner (GP) when needed. However, participants in a control group should receive the same number of sessions, care and attention as participants in the treatment group to adequately correct for placebo and other confounding factors. Also, patients have

signed up to a trial expecting to get some form of treatment in return for taking part in it. Yet in a no treatment or waiting list control group, they have to attend a number of assessments without any direct benefit for themselves. These patients will be disappointed that they have been denied treatment benefits they anticipated from participation in a study. Assignment to no treatment or waiting list may strengthen participants' beliefs that they will not improve, thereby reducing the chance of spontaneous improvement. Participants randomised to these two control conditions may improve less than would be expected compared to participants not enrolled in a trial. Consequently, subjective baseline–follow-up differences cannot be assumed to be the natural history of what would have occurred in the absence of patients enrolling in the study. Therefore, using waitlist or no treatment control conditions can lead to the overestimation of the effectiveness of a treatment (Mohr et al., 2009).

A Cochrane review of non-pharmacological interventions for functional syndromes, including CFS (14 of the 21 RCTs used CBT) (Van Dessel et al., 2014), and a recent meta-analysis of 33 RCTs of mindfulness-based interventions (Dunning et al., 2018), both not only concluded that most benefits disappear when an active control group is used, instead of a waitlist or no treatment control group, but also that there's evidence too that the remaining effects are inflated by bias and multiple methodological concerns, including high drop-out rates and selective biases in sampling. These concerns apply to the studies in the Cochrane CBT review too. It also suggests that studies in the review with a waitlist or no treatment control group might have suffered from similar problems and concluded erroneously that their treatment was effective. A Cochrane review of CBT for schizophrenia by Jones et al. (2018) concluded that

> Considering the maturity of trials in this area, the quality of evidence available is embarrassingly low. The veracity of the findings of the trials is threatened by biases, uncollaborative working, and poor reporting. These issues will have led to much research waste – of funding, and opportunity for researchers, carers, and recipients of care.

These issues also apply to the studies in the Cochrane CBT review and the studies that were still ongoing – and have been published since – and the two studies that have been used by White et al. (2011) in support of their own recovery claims.

5. Problems with subjective outcomes in non-blinded trials

All trials in the review were by definition non-blinded, yet the review used one subjective primary outcome, fatigue. Patient self-report is an unreliable measure (Wechsler et al., 2011). Lack of patient blinding combined with self-reporting of outcomes leads to pronounced bias as patients become prone to outside influences leading to the erroneous inference of efficacy in its absence, thus making subjectively assessed outcomes unreliable (Hróbjartsson et al., 2014; Lilienfeld et al., 2014).

Low correlation between objective and subjective activity measurements (Scheeres et al., 2009) is not only confined to the chronically ill but is also present in the healthy population (Van den Berg-Emons et al., 2011). In non-blinded studies, self-report measures are highly vulnerable to response bias, the size of which is not trivial. No such inflation was observed when objective outcome measures were used (Hróbjartsson et al., 2014). For patients with CFS, there is a particular problem with subjective outcomes as they may feel better able to cope with daily activities, because they have reduced their expectations of what they should achieve, rather than because they have made any recovery as a result of the intervention (Whiting et al., 2001). Also, the self-report measure scores for depression and anxiety only have meaning if the person actually has these disorders. It is important not to automatically pathologise it by using self-report measures. It may simply be that they suffer from inevitable psychological distress which can happen with any serious long-term physical condition made worse by the way patients – with a poorly understood disease like CFS – are treated by society and the medical profession (Blease et al., 2017).

The unreliability of subjective outcomes in non-blinded trials is illustrated by the following examples. In Jason et al. (2007), there was a large difference in SF-36 physical functioning scores at baseline between the relaxation group (RELAX) and the anaerobic activity group (ACT), 53.77 to 39.17. However, the objective physical functioning scores (6MWT) at baseline were almost identical: 1335 (ACT) versus 1317 (RELAX). In the PACE trial (White et al., 2011), there was a subjective *improvement* of physical functioning after CBT of 18.2 per cent (7.1/39.0; SF-36 physical functioning) compared to Specialist Medical Care (SMC), yet the step test (fitness) and 6MWT showed that objectively there was *no difference*. Moss-Morris et al. (2005) reported that subjective physical functioning scores after GET *improved* by 30 per cent (15.95/53.1), yet objectively they *deteriorated* by 15 per cent (4.78/31.99; VO2 peak). The review itself did not acknowledge the unreliability of subjective outcomes in unblinded trials (Price et al., 2008). The only way to correct for this unreliability is by using well-designed control groups and objective primary outcomes (Edwards, 2017; Lilienfeld et al., 2014). The majority of studies (Deale et al., 1996; Huibers et al., 2004; Jason et al., 2007; Lloyd et al., 1993; O'Dowd et al., 2006; Prins et al., 2001; Sharpe et al., 1996; Whitehead and Campion, 2002) used objective outcomes. So it would have been possible for the Cochrane review to have used them. Any conclusion that the intervention was effective must be

seen as unreliable because of the use of subjective outcomes in non-blinded studies.

6. Limitations of the Chalder Fatigue Scale questionnaire

The primary outcome of the review was fatigue using the Chalder Fatigue Scale (ChalderFS) (Price et al., 2008), used by six of the 11 studies in the review that were published in peer-reviewed medical journals (Deale et al., 1997; Ridsdale et al., 2004; Ridsdale et al., 2001; Whitehead and Campion, 2002; Surawy et al., 2005; O'Dowd et al., 2006), or any other validated fatigue scale.

Ten flaws have been identified with the use of the ChalderFS:

1. It does not provide a comprehensive reflection of fatigue-related severity, symptomology or functional disability in CFS (Haywood et al., 2011), as it was developed by mental health professionals and many questions are geared towards depression and not CFS (Chalder et al., 1993).
2. The ceiling effect means that a maximum score at baseline cannot increase even if there is deterioration during the trial. As a consequence, for example, if a participant deteriorated during the trial on eight items and improved on three, the score should reflect a deterioration of five points. However, if he had scored the maximum at baseline, then since eight scores cannot get worse and three scores have improved, the ChalderFS would classify the participant, who has *deteriorated* by five points, as *improved* by three. Analysis of the use of the ChalderFS in CFS patients, who were well enough to attend an outpatient clinic, found high rates of maximal scoring (Morriss et al., 1998). The ceiling effect will therefore be a problem for those patients and an even bigger problem for studies that include the more severely affected.
3. The Scale has been found to be unreliable in distinguishing between healthy controls and fatigue. In a trial of CBT for patients with multiple sclerosis (MS), it was found that after treatment, fatigued MS patients had less fatigue than healthy controls (Van Kessel et al., 2008).
4. Few items on the ChalderFS appear clearly related to fatigue and there is a focus on change in fatigue, rather than intensity (Wilshire et al., 2018a).
5. An accurate comparison relies on the ability to remember pre-illness fatigue, which may be difficult for respondents who have been ill for a long period. They might also wrongly interpret usual state as one of illness instead of one of pre-illness (O'Dowd et al., 2006).
6. The ChalderFS is unable to distinguish between CFS and primary depression (O'Dowd et al., 2006).
7. The items comprising the mental fatigue subscale describe cognitive difficulties rather than mental fatigue, which may not be the same thing (O'Dowd et al., 2006).
8. The ChalderFS fails to interpret the physical lack of energy as a valid cause for reduced cognitive functioning (Roberts, 2018).
9. The ChalderFS index score has limited evidence of test–retest reliability (Haywood et al., 2011).
10. Changing the way of scoring from bimodal to Likert can lead to different outcomes, making the ChalderFS unreliable. This was brought to the attention by the FINE trial. Using bimodal scoring in their original publication, they reported that CBT had no effect on fatigue scores (Wearden et al., 2010b). After publication, they changed the scoring to Likert and now suddenly there was a small but statistically significant improvement (Wearden and Emsley, 2013).

The limitations of the ChalderFS to measure subjective fatigue, as used by the majority of studies in the review, further casts doubt on the trials' and the reviews' conclusions.

7. Concerns over missing data and participant dropouts

The percentage of dropouts/missing data in a number of trials was high and these losses are unlikely to be random (Wilshire, 2017). Drop-out rate may be an important indicator of the acceptability of an intervention. High drop-out rates may indicate that the treatment may accommodate only a very specific group of participants, which will again limit the generalizability of the findings. Where drop-out rates are higher in the intervention group than in the control group, it may be the case that there is something about the intervention that trial participants find unacceptable (Whiting et al., 2001). This was most notable in Prins et al. (2001) with a drop-out rate of 40.9 per cent (CBT) and 23.1 per cent (no treatment). The drop-out rate in some of the other trials was the following: 25 per cent (no significant differences between groups) (Jason et al., 2007) and 31 per cent (CBT) and 36 per cent (counselling) (Ridsdale et al., 2001). There were missing cognitive tests data for 28.9 per cent (15/52, CBT) versus 13.7 per cent (7/51, no treatment) (O'Dowd et al., 2006); missing data for 30.8 per cent (8/26, CBT) and 28.2 per cent (11/39, no treatment) (Whitehead and Campion, 2002) and 28.6 per cent (18/63, CBT) and 40 per cent (24/60, GET) (Ridsdale et al., 2004). Participants who do not respond to treatment or are negatively affected by it are more likely to drop out or be lost to follow-up (Lilienfeld et al., 2014). These dropouts and missing data add further doubts about the reliability of the review's findings.

8. Response-shift bias

CBT for CFS is different from normal CBT, as it aims to cure patients' false illness beliefs and fear avoidance of exercise and activity, by modifying their beliefs and perception of their symptoms. Response-shift bias occurs when an intervention leads individuals to change their evaluation standard with regard to the dimension measured, leading the therapist (and often also the patient) to conclude erroneously that the treatment has worked. The only way to correct for this is by using well-designed control groups and objective outcomes (Lilienfeld et al., 2014).

9. Selection bias

More than half of the studies in the Cochrane review failed to describe randomisation procedures, thus making it impossible to assess the extent to which selection bias may have occurred (Laws, 2017).

10. Participant treatment investment bias

Another important cause of bias in non-blinded trials that use subjective outcomes is effort justification, where patients investing substantial time, energy and effort in an intervention often feel a psychological need to justify this commitment. There is also a tendency for participants to report improvement in accord with what they believe to be the therapist's/researcher's hypothesis. The only way to correct for this is by using well-designed control groups and objective outcomes (Lilienfeld et al., 2014).

11. Concerns about psychiatric comorbidities in CBT studies

A substantial percentage of participants in the studies suffered from comorbid psychiatric disorders as can be seen in Table 1. For example, 38.5 per cent (Deale et al., 1997), 48 per cent (O'Dowd et al., 2006), 58 per cent (Ridsdale et al., 2001), up to 77 per cent in Sharpe et al. (1996) and 74 per cent in Lloyd et al. (1993) had a comorbid depression and 3 per cent a comorbid anxiety disorder. This is of particular concern in CBT studies for CFS, as a meta-analysis by Tolin (2012) found that CBT is the most effective treatment for depression and anxiety disorders. Also, the presence of a medical or psychiatric condition that may explain the chronic fatigue state excludes the classification as CFS in research studies because overlapping pathophysiology may confound findings specific to CFS (Reeves et al., 2003). Finally, none of the CFS studies determined the proportion of people who were no longer suffering from their psychiatric comorbid disorder after treatment with CBT.

12. Problems with the quality of patient-reported outcome measures

A systematic review of patient-reported outcome measures (PROMs) in CFS found poor quality of the reviewed PROMs which included the ChalderFS. Combined with the failure to measure outcomes relevant to patients, this suggests that high-quality and relevant information about treatment effect is lacking (Haywood et al., 2011).

Analysis of the trials

Barrett (1992)

This unpublished study did not go through peer review and is not available on the Internet.

Lloyd et al. (1993)

The authors found that neither immunological therapy (dialyzable leukocyte extract) nor CBT (alone or in combination) provided greater benefit than the nonspecific treatment regimens. The graded exercise regime in the CBT group programme did not lead to significantly increased levels of non-sedentary physical activity. Lloyd et al. (1993) concluded that their results do not support the hypothesis that CBT is an adequate treatment for CFS.

Sharpe (1993), published as Sharpe et al. (1996)

Sharpe (1993) is a literature review of the role of psychological (cognitive) and behavioural therapies in CFS. In it, a RCT of CBT – then in progress – was described which was published as Sharpe et al. (1996). This study had a no treatment control group and 35 per cent of participants did not have any impairments of daily activities at trial entry (a Karnofsky score of 80 or more). The inclusion criteria included meeting the Oxford criteria and a Karnofsky score <80, indicating impairment of daily activities; 48.8 per cent (60/123) of participants screened were selected; 67 per cent had a comorbid anxiety or depression and 10 per cent a somatisation disorder.

The groups were poorly matched: 16 one-hour individual treatment sessions (CBT) and 0 (no treatment control group; both groups could see their own GP if needed). There were twice as many men in the CBT group (40%, 12/30) as in the control group (23%, 7/30). After 5 months of being stimulated to gradually and consistently increase their activity levels, they had improved their 6MWT by 9.9 per cent (42/424) compared to no treatment. At 12 months, this had improved by another 3.1 per cent (13/424). Exercise was not part of the control treatment (no treatment). At 12 months, depression improved by 29.9 per cent (2.0/6.7) more after CBT compared to no treatment. The improvement in 6MWT might therefore reflect an improvement in their depression, present in 67 per cent of participants. There are also other

Table 1. Summary of the findings of the trials in the Cochrane review.

Studies	Treatment	No. of participants	Selection criteria	Control group	Psychiatric comorbidity	Objective outcomes/work	Quality of life	Missing data/dropouts
Barrett, 1992	Unpublished study							
Deale et al., 1997	CBT	60	Oxford	Relaxation; poorly matched	38.5%	No differences (at 5 years) between groups (employment status, physical functioning, fatigue, general health, meeting CFS criteria)	Not used	10% CBT, 13% (relax) dropouts at 6 months; 16.7% CBT, 6.7% relax at 5 year (non-completers)
Huibers et al., 2004	CBT delivered by GPs	66 CFS patients (fatigue study: $n = 151$)	Fukuda	No treatment	Unclear	At 4 months: 50% (CBT) and 61% (NT); at 12 months: 59% and 65% resumed work; Clinical recovery at 12 months: 33% and 44%. Actometer results not published	Not used	Did not complete: 33% CBT 0% no treatment
Jason et al., 2007	CBT	114 (54% self-selected)	Fukuda	Relaxation	38.6% depression	+6.1% (6MWT)	−5% CBT vs relax	25% (dropouts)
Lloyd et al., 1993	Immunological and CBT	90	Lloyd	Non-specific	74% depression 3% anxiety	No improvement (T-lymphocyte count + activity diary)	Not used	No information
O'Dowd et al., 2006	Group CBT	153	Fukuda	No treatment; poorly matched	48% depression; 48% on SSRIs, 11% on benzodiazepines	No improvement (neurocognitive performances, work status or shuttles walked); walking speed improved more (CBT)	No improvement	Missing cognitive test data: 28.9% CBT and 13.7% no treatment
Prins et al., 2001	CBT	278	Oxford	No treatment; not evenly matched	No information provided	No improvement (actometer, work status, neuropsychological tests)	No improvement	40.9% CBT and 23.1% no treatment (dropouts)

(Continued)

Table 1. (Continued)

Studies	Treatment	No. of participants	Selection criteria	Control group	Psychiatric comorbidity	Objective outcomes/work	Quality of life	Missing data/dropouts
Ridsdale et al., 2001	CBT	45 CFS patients (fatigue study, $n = 160$)	Fukuda	Counselling	58% anxiety and/or depression	Not used	Not used	36% counselling and 31% CBT (dropouts)
Ridsdale et al., 2004	CBT and GET	36 CFS patients (fatigue study: $n = 123$)	Fukuda	Post hoc added non-randomised prospective no treatment control group; poorly matched	History of anxiety or depression: 60% CBT and 57% GET	Step test results not published	Not used	28.6% CBT and 40% GET (missing data)
Russell et al., 2001	Unpublished study							
Sharpe et al., 1996	CBT	60 (35% no impairments of daily activity at trial entry)	Oxford	No treatment; poorly matched	67% depression or anxiety; 10% somatisation disorder	6MWT: 9.9% (5 months), additional 3.1% (12 months); depression improved by 29.9% (12 months)	Not used	No information
Stevens et al., 1999b	Unpublished study							
Strang, 2002	CBT (unpublished non-peer-reviewed study)	51	Only abstract available on the Internet	Waitlist	Unclear	No changes in fatigue or functional impairment	Unclear	25.5% (missing data)
Surawy et al., 2005	Mindfulness	41 (18 + 12 + 11)	Oxford	Study 1: waitlist; study 2 + 3: non-randomised no control group	No information provided	Study 1 + 2: no improvement (fatigue and physical functioning)	Not used	25% study 2 and 18.2% study 3 (dropouts)
Whitehead and Campion, 2002	CBT delivered by GPs	65	Fukuda	No treatment; not evenly matched	No information provided	Patients remained highly disabled over the 12-month study period	Not used	30.8% CBT and 28.2% no treatment (missing data)

CBT: cognitive-behavioural therapy; CFS: chronic fatigue syndrome; GP: general practitioner; NT: no treatment; 6MWT: 6-minute walk test; GET: graded exercise therapy.

reasons to consider this benefit with caution, as discussed later during the analysis of the objective outcomes.

One cannot safely conclude that CBT is an effective treatment, in view of the poorly matched groups, using the Oxford criteria, a no treatment control group, not excluding the 35 per cent of participants who had no impairments of daily activities at trial entry and using CBT in a trial where 67 per cent of participants had a comorbid depression.

Deale et al. (1996), RCT published as Deale et al. (1997)

This study used the Oxford criteria; 38.5 per cent in the study had a comorbid psychiatric disorder. The groups were poorly matched: there was a difference in mean ages (31 years CBT, 38 years relaxation), illness duration (3.4 years CBT, 4.6 years relaxation), unemployed (63% and 77%), disability benefits (53% and 67%), past psychiatric disorder (30% and 13%) and attribution of symptoms to physical illness (57% and 73%) at baseline. This also indicates that 43 per cent of participants in the CBT group knew that they had a psychiatric disorder in a psychotherapy study. CBT contained a graded exercise programme; the relaxation control group did not receive any exercise.

The General Health Questionnaire scores (0–12) at baseline were 6.2 and 6.0 indicating psychological caseness according to the authors (scores of 4 or more). The scores had improved by 14.5 per cent (0.9/6.2) after CBT compared to relaxation and there was no psychological caseness anymore after CBT (3.4), contrary to relaxation (4.3). The study concluded that at 5-year follow-up, CBT can produce some lasting benefits but is not a cure for CFS. Yet since completing treatment and before the 5-year follow-up, 56 per cent (CBT) and 57 per cent (relaxation) received further treatment for CFS (Deale et al., 2001). There were also no statistically significant differences at 5-year follow-up between the groups in physical functioning score (more than 83; $p=0.27$), the Fatigue Questionnaire (less than 4, $p=1.00$), the General Health Questionnaire ($p=0.58$), no longer meeting the Oxford CFS diagnostic criteria ($p=0.42$) and employment status ($p=0.28$) (Deale et al., 2001).

The authors acknowledge that the study has its limitations. These include the use of a single therapist who administered both treatments (Deale et al., 1997) and self-rated outcome measures. They also acknowledged that it is difficult to draw firm conclusions about the effect of CBT given that many patients received further treatment during follow-up (Deale et al., 2001).

In conclusion, one cannot safely conclude that CBT is effective in view of the above named problems, the fact that there is no difference in fatigue, physical functioning, general health or no longer meeting the Oxford criteria at long-term follow-up and the lack of objective improvement (employment status).

King (1999), published as Ridsdale et al. (2001)

This was a trial of 160 patients with chronic fatigue; 28 per cent met research criteria for CFS (Fukuda; post hoc analysis), 58 per cent had a history of anxiety and/or depression and 21 per cent were on antidepressants at baseline. Drop-out rate was 31 per cent (CBT) and 36 per cent (counselling). Objective outcomes were not used. The authors concluded that CBT was not more effective than counselling.

Stevens et al. (1999b)

This unpublished study did not go through peer review and is not available on the Internet.

Prins et al. (2001)

This study states that it used the Fukuda criteria with the exception of the criterion, requiring four of eight additional symptoms; therefore, it used the Oxford and not the Fukuda criteria. The groups were not evenly matched. The number of sessions was as follows: 16 sessions of 1 hour (CBT), 11 sessions of 1.5 hour (guided support by a social worker) and 0 (no treatment); sickness impact profile (SIP) total score 1755 (CBT) and 1859 (no treatment); hours worked 16.3 and 13.5; quality of life scores 46 and 40 (higher score indicating better quality of life) for CBT and the no treatment group, respectively. These differences indicate that the no treatment group might have been more disabled.

In all, 59.1 per cent (55/93, CBT) and 76.9 per cent (70/91, no treatment) completed the study. Drop-out rate was therefore: 40.9 per cent (CBT) and 23.1 per cent (no treatment). The number of hours worked suggests that participants had mild CFS. CBT was not more effective after 14 months than no treatment for psychological well-being and focusing on bodily symptoms, measured by the symptom checklist 90 ($p=0.1767$), and after 8 and 14 months for the quality of life ($p=0.1878$ and $p=0.4619$) or the number of hours worked ($p=0.3362$ and $p=0.1134$).

The actometer results were not published. Analysis 9 years later by two of the authors of the study showed that CBT did not lead to objective improvement (Wiborg et al., 2010). Analysis of the objective neuropsychological tests (two reaction time tests and a symbol digit modalities task) which were available for 83.8 per cent (233/278) of participants, equally divided over the three groups (78 CBT; 79 support group; 76 no treatment), by three of the authors, also showed that CBT did not lead to objective improvement (Knoop et al., 2007b).

One cannot safely conclude that CBT is more effective than guided support by a social worker or than the natural

course (no treatment), in view of the above named problems and the lack of objective improvements.

Russell et al. (2001)

This unpublished study did not go through peer review and is not available on the Internet.

Strang (2002)

This is an unpublished study that did not go through peer review. Its abstract is available on the Internet. Fifty-one CFS patients were randomly assigned to CBT or a wait-list control condition. There were missing data for 25.5 per cent (13/51). CBT was associated with an improvement in mood, but did not lead to changes in fatigue severity, functional impairment, maladaptive cognitions, depression or anxiety.

Whitehead and Campion (2002)

The authors concluded that CBT delivered by a GP, compared to usual care (no scheduled treatment), had no effect on the illness of the patients. Also, patients remained highly disabled over the 12-month study period, despite being 5 years younger (36 CBT, 41 no treatment), and having been ill for a mean of 12-month less (21 CBT, 33 no treatment).

Huibers et al. (2004)

In this fatigue study, CBT was less effective then no treatment in getting people back to work; 56 per cent had unexplained persistent fatigue and 44 per cent met research criteria (Fukuda) for CFS. The objective actometer results were not published. Less patients resumed work in the CBT than in the no treatment group, at 4 months (end of treatment) (50% and 61%) and at 12 months follow-up (59% and 65%). Also, they resumed work slower than in the no treatment group. The percentage who made a clinical recovery at 12 months follow-up was as follows: 33 per cent (CBT) and 43 per cent (no treatment).

The authors concluded that there was no significant difference between CBT delivered by GPs – five to seven 30-minute sessions – and no treatment on primary or secondary outcomes at any point. Also, that the lack of efficacy is likely to result from a disturbance in the interaction between the patient, the doctor and the intervention as treatment was delivered under 'ideal circumstances'. Yet their study showed that doing nothing is more effective in getting people with unexplained persistent fatigue and CFS back to work than CBT.

Ridsdale et al. (2004)

This fatigue study did not publish its objective outcome. Only 29 per cent (36/123) in the two treatment groups had Fukuda defined CFS, and it had a post hoc added, non-randomised prospective no treatment control group in a different time frame. It is not specified how many CFS patients there were in the control group.

The groups were poorly matched: 4.6 sessions – mean number of 45-minute sessions over 12 weeks – (both treatment groups: CBT and GET), 0 sessions (no treatment control group; they were only given a CBT booklet for fatigue self-management); number of participants: 63 (CBT), 60 (GET) and 40 (no treatment). A large percentage of participants were working at baseline: 60 per cent (CBT) and 73 per cent (GET), indicating that patients were higher functioning; 50 per cent in both treatment groups had a history of past psychiatric referral compared to only 30 per cent in the control group; 60 per cent (CBT) and 57 per cent (GET) had a history of anxiety/depression. There were missing data (data were analysed 'per-protocol', including only patients who completed six sessions of therapy) for 28.6 per cent (18/63, CBT) and 40 per cent (24/60, GET).

The authors acknowledged that one of the limitations of their trial was a non-randomised control group done in a different time frame. Also, it appears that there were no CFS patients in the control group. Ridsdale et al. (2004) concluded that there was a significant difference in groups receiving therapy and no treatment in percentage recovered but that GET is not superior to CBT. The only recovery criterion however was no longer fulfilling the Chalder case criteria for fatigue. It is not clear what this score was, as patients were recruited using binary scoring (0–11) with a cut-off score of 4 or more; yet during the trial, a Likert score was used (0–33). A binary score of 3 or less equals to a Likert score of 6 to 9 or less. The mean fatigue scores at 8 months were the following: 14.8 (CBT), 15.3 (GET) and 17.9 (no treatment).

The objective outcome (step test) was not published. This jeopardises the validity of a study (Heneghan et al., 2018). In the largest CBT trial so far (White et al., 2011), the step test showed that CBT and GET did not lead to objective improvement. A recent reanalysis of the Cochrane GET review showed that GET does not lead to clinically significant objective improvement (Vink and Vink-Niese, 2018b). The authors acknowledge that the subgroup analysis of patients with CFS was too small to provide power. In view of the aforementioned problems, one cannot safely conclude that CBT is an effective treatment.

Surawy et al. (2005)

This study that used the Oxford criteria consisted of three small studies of 8 weeks of mindfulness. In study 1 ($n = 18$), 40.9 per cent (18/44) who were asked, agreed to take part. Mindfulness did not lead to a statistically significant change in subjective fatigue ($p = 0.08$) and physical functioning ($p = 0.58$), compared to the waiting list control group.

In study 2, only 44 per cent (12/27) of patients who had been on the waiting list in study 1 agreed to take part; of

those, 25 per cent (3/12) dropped out. Subjective fatigue ($p=0.08$) and physical functioning ($p=0.58$) did not improve.

In study 3, this was 36.7 per cent (11/30) and 18.2 per cent (2/11). Both studies (2 + 3) were not randomised, did not have a control group and were therefore not properly controlled. None of the studies used objective outcomes. Studies 1 and 2 showed that mindfulness is not effective. One cannot safely conclude anything about the effectiveness of mindfulness in the other study, because of the above-mentioned flaws.

O'Dowd (2000) (trial register), RCT published as O'Dowd et al. (2006)

Group CBT is not more effective than no treatment; 37 per cent (CBT) – 30 per cent, three groups together – already had a SF-36 physical functioning score within the normal range at trial entry. Its groups were poorly matched: 8 meetings, each lasting 2 hours (both the CBT and education and support (EAS) groups), 0 (no treatment group; labelled as specialist medical care; participants could see their GP if needed). There were no significant differences between groups in number of GP consultations: 5.8 (CBT), 6.0 (EAS) and 6.5 (no treatment). There were almost twice as many men in the CBT group in comparison with EAS and no treatment (24, 12 and 15, respectively). At baseline, 27 per cent had a HADS (Hospital Anxiety and Depression Scale) depression score of more than 11, indicating depression or anxiety 'caseness' in the CBT group, but only 12 per cent in the no treatment group.

In total, 89 per cent had been working prior to their illness; for 70 per cent, CFS prevented them from continuing to do that. This suggests that 19 per cent were still working and therefore higher functioning; 52 per cent had a SF-36 mental health score within the normal range at trial entry, indicating that 48 per cent of participants had a comorbid mental health problem. This was confirmed by the high proportion of participants who used Selective Serotonin Reuptake Inhibitors (SSRIs, 48%) and benzodiazepines (11%) during the 12 months of the trial. At 12 months, there were missing cognitive tests data for 28.9 per cent (15/52, CBT) and 13.7 per cent (7/51, no treatment) and the physical functioning score was in the normal range for 46 per cent (CBT) and 44 per cent (no treatment); also, 32 per cent (CBT) and 49 per cent (no treatment) showed at least a 15 per cent increase in physical function.

There were no significant differences in the neurocognitive performances (simple reaction time, repeated digits detection data and mood scores; $p \geqslant 0.21$ for all tests) or the numbers of shuttles walked between the three groups ($p=0.16$). Therefore, fitness did not improve. Walking speed improved more in the CBT group. However, there are reasons to consider this benefit with caution, as discussed later in the analysis of the objective outcomes. Also, the authors themselves wondered about the clinical relevance of it. The authors concluded that group CBT did not significantly improve cognitive function, healthcare utility measures, quality of life or employment status. O'Dowd et al. (2006) showed that group CBT is not more effective than no treatment.

Jason et al. (2007)

This trial had a biased sample: 54 per cent of participants had contacted the university, so were self-selected. At baseline, only 24.6 per cent were on disability benefits. This suggests that the percentage of higher functioning CFS patients was 75.4 per cent. Drop-out rate was 25 per cent (no significant differences between groups). The objective 6MWT results after CBT improved by 6.1 per cent (196.25/1346.35 – 111.55/1317.78) compared to relaxation. The core symptom of the disease – PEM – improved by 4.2 per cent (26.82/77.5 – 20.63/67.88) after CBT compared to relaxation.

The Quality of Life scores after CBT improved by 5.0 per cent less (2.96/66.14 – 6.25/65.75) than after relaxation. These scores at the 12-month follow-up were 69.1 (CBT) and 72.0 (RELAX) (16–112; higher scores indicating higher quality-of-life). This was still worse than fibromyalgia (70), chronic obstructive pulmonary disease (COPD), psoriasis and urinary incontinence (82), rheumatoid arthritis (83), systemic lupus erythematosus (84), osteoarthritis (87) and young adults with juvenile rheumatoid arthritis (92) (Burckhardt and Anderson, 2003).

The RELAX group performed better than the CBT group in regard to improvement for sore throat, muscle pain, unrefreshing sleep and headaches. The authors themselves concluded that there were few significant differences for employment or the participant and clinician ratings among the four conditions at 12-month follow-up. Also, that no treatment strategy was clearly superior to another treatment strategy in all areas and that the limitations of the study included the absence of a control group and that more participants were working and less were on disability benefits at baseline than in other studies.

Review of the objective outcomes

The majority of studies (Deale et al., 1997; Huibers et al., 2004; Jason et al., 2007; Lloyd et al., 1993; O'Dowd et al., 2006; Prins et al., 2001; Sharpe et al., 1996; Whitehead and Campion, 2002) used objective outcomes. Sharpe et al. (1996) showed an improvement of 9.9 per cent in the 6MWT at 5 months (end of treatment) and an additional 3.1 per cent at 12 months, of the intervention group compared to no treatment. In Jason et al. (2007), the 6MWT results after CBT improved by 6.1 per cent compared to relaxation. In both studies, exercise was an element of CBT but not of the control treatment (no treatment and relaxation, respectively).

Moreover, there are also other reasons to consider these benefits with caution. Sharpe et al. (1996) used the Oxford criteria, it had a high percentage of participants with a psychiatric disorder and included 35 per cent of high-functioning participants who did not have any impairment of daily activities at trial entry. The no treatment control group was poorly matched and at 12 months, depression improved by 29.9 per cent (2.0/6.7) more after CBT than after no treatment, which might account for the small objective improvement. Jason et al. (2007) included a high percentage of high-functioning participants, only 24.6 were on disability benefits at baseline, and it had a drop-out rate of 25 per cent for the whole study (no statistical differences between groups).

The learning effect in the 6MWT – better performing on the second test because of familiarity with the test – in 761 patients with severe emphysema (mean age of 67, 10% requiring oxygen at rest and 77% during the test) was 7 per cent, and in patients with hypoxemia from COPD and restrictive lung diseases, it was 14.9 per cent (Sciurba et al., 2003). Stevens et al. (1999a) performed three 6MWTs in COPD patients to see how much patients would improve by simply repeating the test on consecutive days. They found a learning effect (a mean increase) of 10 per cent on the second test and an additional 3 per cent on the third test. This is the same improvement as reported by Sharpe et al. (1996) in their tests. Also, the small improvements in both Sharpe et al. (1996) and Jason et al. (2007) are in stark contrast with patients with stable chronic heart failure, who improved their 6MWT results by 65 per cent after only 3 weeks of exercising (Meyer et al., 1997). Finally, a major criterion for defining CFS is a reduction in physical capacity of at least 50 per cent compared to pre-illness levels (Carruthers et al., 2011; Fukuda et al., 1994; Holmes et al., 1988); so the small improvements in both studies would still leave patients considerably worse off than before the illness.

CBT, which included a graded exercise regime, did not lead to significantly increased levels of non-sedentary physical activity (standardised diaries of daily activities) or improvement of T-lymphocyte count in Lloyd et al. (1993). In Deale et al. (1997), there was no significant difference in employment status at long-term follow-up between CBT and relaxation ($p = 0.28$), despite using the Oxford criteria and a poorly matched control group.

In O'Dowd et al. (2006), there were no significant differences in the objective neurocognitive performance (simple reaction time, repeated digits detection data and mood scores; $p \geqslant 0.21$ for all tests) or the numbers of shuttles walked between the three groups ($p = 0.16$). Therefore, fitness did not improve. Walking speed improved more in the CBT group. However, there are reasons to consider this benefit with caution. A high percentage of participants had a psychiatric disorders; 19 per cent of participants were still working; in the CBT group, 37 per cent had a physical functioning score within the normal range at trial entry and it included twice as many men compared to the control group in a trial where exercise was an element of CBT but not of the control group (no treatment). There were also twice as many missing data in the CBT (28.9%) then in the no treatment group (13.7%). At 12 months, the physical functioning score was in the normal range for 46 per cent (CBT) and 44 per cent (no treatment); 32 per cent (CBT) and 49 per cent (no treatment) showed at least a 15 per cent increase in physical function. The authors themselves concluded that 'it is not clear whether the improvement observed [in walking speed] is clinically significant' (p. 43). Finally, after CBT, participants walked 28.4 (10 m) shuttles; as noted by the authors, healthy subjects (mixed gender and age) walked a mean of 67 (10 m) shuttles (O'Dowd et al., 2006).

Prins et al. (2001) did not publish its actometer results. Analysis 9 years later, by two of the authors of the study, showed that CBT did not lead to objective improvement (Wiborg et al., 2010). Analysis of the objective neuropsychological tests (two reaction time tests and a symbol digit modalities task) which were available for 83.8 per cent of participants, by three of the authors, also showed that CBT did not lead to objective improvement (Knoop et al., 2007b). Moreover, CBT was not more effective after 8 and 14 months than no treatment, for the number of hours worked ($p = 0.3362$ and $p = 0.1134$).

Huibers et al. (2004) did not publish its actometer results. Analysis of the results from three other studies, that also did not publish those results, by proponents of the biopsychosocial model, showed that CBT did not lead to objective improvement (Wiborg et al., 2010). Huibers et al. (2004) showed that CBT is less effective in getting people back to work than no treatment. Also, they resumed work slower than in the no treatment group.

Ridsdale et al. (2004) did not publish its objective outcome (step test), just like Prins et al. (2001) and Huibers et al. (2004). It is unlikely that they would not have published these results, if they had supported their hypothesis that CBT is effective. Stordeur et al. (2008) found no objective improvements after CBT and GET (VO2max) in the Belgian CFS knowledge centres.

Quality of life

Three trials from the review recorded quality-of-life scores. CBT was not more effective in Prins et al. (2001), than no treatment for the quality of life. In Jason et al. (2007), quality-of-life scores improved 5 per cent more after relaxation than after CBT. Group CBT did not bring about improvement in quality of life in O'Dowd et al. (2006).

Discussion

Treatment guidelines, particularly the promotion of CBT as a treatment for ME/CFS, have been influenced by Cochrane

reviews such as Price et al. (2008). In this article, we reanalysed that Cochrane review and the trials in it. Four of the 15 studies were unpublished and did not go through peer review. Only the abstract is available on the Internet for one of those. Three studies were fatigue studies, in which only a minority of participants had CFS, and one was a mindfulness study. PEM, the main characteristic of CFS, was only compulsory for diagnosis in one of the 15 studies (Lloyd et al., 1993). This study concluded that CBT is not an adequate treatment for CFS. The CBT trials reviewed here are inherently biased as patients have to be able to attend outpatient clinics which the more severely disabled, those who are home or bedbound – 25 per cent according to most estimates (Pendergrast et al., 2016) – are unable to do. Therefore, findings from RCTs are not generalisable to the wider CFS population. The only trial that looked at the more severely affected was the FINE trial by Wearden et al. (2010b; see Appendix 1 for a more detailed analysis). It was published 2 years after the Cochrane review and found that CBT and GET do not lead to objective improvement (step test) in the more severely affected.

Our analysis of the RCTs included in the Cochrane CBT review identified 12 areas of concern. These included potentially selecting patients who do not have the disease. A failure to exclude patients with comorbid depression and anxiety, even though CBT is the most effective treatment for both, as a meta-analysis by Tolin (2012) found. One of the other important problems of the trials and the review itself was the use of subjective primary outcomes even though in non-blinded trials, self-reported outcomes can produce highly inflated estimates of treatment-related benefits contrary to objective outcome measures (Wilshire et al., 2018b). Moreover, it is unclear why the trials relied on subjective primary outcomes when the basis of CBT for CFS is that patients suffer from false illness beliefs and they do not know how to interpret their symptoms correctly. The flaws in the review and the trials, as discussed at the beginning of this article, all created a bias in favour of the intervention. It was therefore crucial to demonstrate accompanying improvement on more objective measures yet these did not yield significant treatment effects. Most notably, treatment did not affect fitness.

Our analysis shows that CBT did not improve the quality of life scores either, in a disease where these scores are much lower than in the general population and the lowest compared to 20 other chronic illnesses which included stroke, lung cancer and MS (Falk Hvidberg et al., 2015). Nothing has changed in that respect compared to the 1996 health status report (Komaroff et al., 1996) despite the widespread use and promotion of CBT and GET as effective treatments, which provides indirect evidence of their inefficacy.

The Cochrane review by Price et al. (2008) found that the added benefit of CBT over no-care (usual care) on fatigue is 14 per cent. Therefore, just one out of seven report more improvement of their fatigue from CBT than from doing nothing. Friedberg and Krupp (1994), which was excluded from the Cochrane review by Price et al. (2008), found that CBT helped to reduce the symptoms of depression, stress and fatigue severity in depressed but not in non-depressed CFS patients. It might therefore well be that the small improvement in fatigue in a minority of patients simply reflects an improvement of their comorbid depression. The small added subjective benefit in fatigue could also be explained by a range of biases that dog CBT RCTs as discussed at the beginning of this article. Price et al. (2008) also concluded that there is little evidence for long-term benefits using CBT.

Real-world application of CBT in NHS CFS clinics shows equally poor results. Moreover, patients who had been treated with CBT and GET in these clinics had less improvement in fatigue at 12 months than those who had been offered activity management (Crawley et al., 2013). This is in line with the outcomes of patient surveys which have repeatedly shown that rest and activity management (pacing) are the most helpful, with CBT and GET among the least effective therapies (Kirke, 2017). Just 3 per cent of CFS patients report no longer having CFS after NHS treatment; 2–5 years after the initial assessment, this was 5.7 per cent (Collin and Crawley, 2017), which is essentially the same as the naturally occurring recovery rate of 5 per cent (Cairns and Hotopf, 2005).

The impact of CFS results in disruptions to productivity and meaningful occupation, which is often not the case with other conditions (Roberts, 2018). An influential systematic review by Cairns and Hotopf concluded in 2005 that because there is increasing evidence for the effectiveness of CBT and GET, that 'Medical retirement should be postponed until a trial of such treatment has been given'. Yet our reanalysis, just like the reanalysis of the Cochrane GET review (Vink and Vink-Niese, 2018b), shows that CBT and GET do not improve the number of hours worked or sickness and disability benefit status. Essentially, these findings do not differ from the evaluation of the effectiveness of CBT and GET in the Belgian CFS knowledge centres (Stordeur et al., 2008) or in the NHS CFS clinics (Collin and Crawley, 2017). In other words, undergoing these treatments should not be a requirement to be eligible for medical retirement.

The Cochrane review by Price et al. (2008) did not use objective outcome measures even though they were used by the majority of trials. Our analysis of the objective outcomes of these trials provides sufficient evidence to conclude that CBT is not an effective treatment for ME/CFS. This is confirmed by the lack of objective improvement – VO2max – after CBT in real life in the Belgian CFS knowledge centres (Stordeur et al., 2008).

According to the reviewers, there were six trials still in progress which will help to strengthen the evidence base on CBT interventions for CFS. Three of those studies have not

been published, despite being registered in 2002 and 2008. This suggests that their results did not support their hypothesis as negative findings are an important reason for not publishing trial results (Turner et al. 2008). The other three, plus two studies that White et al. (2011) – the biggest CBT trial for CFS ever conducted – used to support their own recovery claims, and one RCT published in 2011, that was excluded from the Cochrane GET review, have also been analysed in this reanalysis (see appendices 1–3). These six studies (Flo and Chalder, 2014, $n=200$; Knoop et al., 2007a, $n=96$; Núñez et al., 2011, $n=120$; Wearden et al., 2010b, $n=296$; White et al., 2011, $n=640$; Wiborg et al., 2015, $n=204$) included 1556 participants, compared to 1008 participants with CFS, in published peer-reviewed studies, in the Cochrane review. There were also many issues with these studies (see Appendices 1 and 2). Objective outcomes were only used by two of these six studies (Wearden et al., 2010b; White et al., 2011). They also showed that CBT does not lead to objective improvement.

The failure of CBT to 'reverse' CFS is perhaps not so surprising when we consider recent exercise physiology studies (Wilshire et al., 2018b). For example, a meta-analysis by Franklin et al. (2018) of 32 studies found that CFS patients have a substantially reduced VO2peak compared to healthy sedentary controls. Therefore, this physiological abnormality cannot be attributed to deconditioning or false illness beliefs. That also implies that the theoretical model upon which CBT for CFS is based is incorrect.

Most studies in Price et al. (2008) failed to report on safety or adverse reactions. Kindlon (2011) and Geraghty et al. (2017b) who pooled patients' surveys ($n=1808$, 5 surveys and $n=3251$, 10 surveys, respectively) found that 20 per cent of respondents reported that CBT had worsened their health. After treatment in NHS CFS clinics, this was 26.8 per cent (18.8% (12.2/65) reporting slight deterioration and 8% reporting much worse or very much worse health) (Collin and Crawley, 2017). Given these considerations, one cannot conclude that CBT is safe either.

Two American government agencies, the CDC (2017) and the AHRQ (Smith et al., 2016), have recently removed (CDC) and downgraded (AHRQ) their recommendations for CBT and GET, because there is insufficient evidence that they are effective. The Dutch Health Council (2018) concluded in March 2018 that CBT and GET are not considered to be adequate treatments for CFS according to general medical standards (in Dutch: 'CGT en GET zijn bij ME/CVS niet te beschouwen als naar algemeen medische maatstaven adequate behandelingen' (p. 6)). The IOM concluded in 2015 that there are no effective treatments for this multisystem disease.

As concluded by Wilshire et al., (2018b),

> If one were to ask, Given the procedures used here, what pattern of results would we expect if these therapies did not produce genuine change? the answer would be, Modest, short-lived changes in self-report behaviour unaccompanied by objectively measurable changes – a pattern much like the one obtained.

The time has come to downgrade CBT to adjunct support-level status and only use it if patients need help coping with a debilitating disease or with a comorbid depression or anxiety disorder.

Conclusion

Seven patients need to be treated for one to report a small, short-lived subjective improvement of fatigue in patients with mild CFS. This is not matched by an objective improvement of physical fitness, disability and sickness benefit status or hours worked. Most studies in the Cochrane review failed to report on safety or adverse reactions. Patient evidence suggests adverse outcome after CBT in 20 per cent of cases. If a trial of a drug or surgical procedure uncovered a similar high rate of adverse outcome, it would be unlikely to be accepted as safe. It is time to downgrade CBT to an adjunct support-level therapy, rather than a treatment for ME/CFS.

Implications for practice

1. For clinicians

If a skilled CBT therapist is available, who acknowledges the severity of this debilitating multisystem disease, then offering CBT to patients as an adjunct support therapy is something to consider. Especially when patients also suffer from a comorbid depression or anxiety disorder or have problems adapting to a life of disability/dependence on others. The findings of this reanalysis do not encourage to instigate treatment programmes of CBT for people with CFS as has been the case so far.

2. For policymakers, occupational health services and illness benefit assessors

This reanalysis shows that CBT does not lead to an improvement of fitness, a reduction of the number of patients on sickness and disability benefits or an improvement of employment status. Forcing patients to undergo this expensive treatment as a requirement to be eligible for illness benefits or medical retirement causes a lot of stress for patients and their families without any benefits to patients and society.

3. For people with CFS

For many patients with CFS, becoming involved in a CBT programme has been compulsory to be eligible for illness benefits or medical retirement. Based on the notion that if

they did not want to be treated with CBT, they were after secondary gains and not motivated to get better. This reanalysis however shows that CBT does not lead to significant improvement in quality of life. Nor does it lead to an improvement in fitness or employment status or reduction of the number of patients receiving sickness and disability benefits. In any event, people with CFS, if offered this therapy, should know that any effect on fatigue is likely to be short-lived and small in degree. Also, that six out of seven patients will undergo the treatment without any benefit and one in five will suffer negative consequences because of it. However, if patients suffer from a comorbid depression or anxiety disorder, or they need help coping with a debilitating illness, then it would be wise to consider help or support from a qualified and knowledgeable psychologist.

Acknowledgements

The authors would like to thank M.V.'s parents for typing out his speech memos. The authors would also like to thank the reviewers for their help in improving the article.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

References

Baraniuk J (2017) Chronic fatigue syndrome prevalence is grossly overestimated using Oxford criteria compared to Centers for Disease Control (Fukuda) criteria in a U.S. population study. *Fatigue: Biomedicine, Health & Behavior* 5(4): 215–230.

Barrett E (1992) *Cognitive-behaviour therapy in the management of chronic fatigue syndrome*. M Psych, University of Sydney, Sydney, NSW, Australia.

Blease C, Carel H and Geraghty K (2017) Epistemic injustice in healthcare encounters: Evidence from chronic fatigue syndrome. *Journal of Medical Ethics* 43(8): 549–557.

Bleijenberg G (2008) The effectiveness of cognitive behavioural therapy in groups for patients with chronic fatigue syndrome (CFS): A randomised controlled study. *Controlled-trials.com*. Available at: http://www.isrctn.com/ISRCTN15823716

Burckhardt CS and Anderson KL (2003) The Quality of Life Scale (QOLS): Reliability, validity, and utilization. *Health Qual Life Outcomes* 1: 60.

Cairns R and Hotopf M (2005) A systematic review describing the prognosis of chronic fatigue syndrome. *Occupational Medicine* 55: 20–31.

Carney C and Jones-Alexander J (2008) Behavioral insomnia therapy with chronic fatigue syndrome. Available at: Clinicaltrials.gov

Carruthers BM, van de Sande MI, De Meirleir KL, et al. (2011) Myalgic encephalomyelitis: International Consensus Criteria. *Journal of Internal Medicine* 270: 327–338.

Centers for Disease Control and Prevention [CDC] (2017) Treatment update ME/CFS. *CDC*, May 30. Available at: https://www.cdc.gov/me-cfs/treatment/index.html (accessed 15 November 2018).

Chalder T, Berelowitz G, Pawlikowska T, et al. (1993) Development of a fatigue scale. *Journal of Psychosomatic Research* 37(2): 147–153.

Collin SM and Crawley E (2017) Specialist treatment of chronic fatigue syndrome/ME: A cohort study among adult patients in England. *BMC Health Serv Res* 17(1):488.

Crawley E, Collin SM, White PD, et al. (2013) Treatment outcome in adults with chronic fatigue syndrome: A prospective study in England based on the CFS/ME National Outcomes Database. *QJM* 106(6): 555–565. Erratum in: *QJM*. DOI: 10.1093/;106(6):567.

David AS (1991) Post viral syndrome and psychiatry. *British Medical Bulletin* 47: 966–988.

Deale A, Chalder T, Marks I, et al. (1996) *Cognitive behaviour therapy for chronic fatigue syndrome: A randomised controlled trial*. London: South Thames Research and Development.

Deale A, Chalder T, Marks I, et al. (1997) Cognitive behavior therapy for chronic fatigue syndrome: A randomized controlled trial. *American Journal of Psychiatry* 154: 408–414.

Deale A, Husain K, Chalder T, et al. (2001) Long-term outcome of cognitive behavior therapy versus relaxation therapy for chronic fatigue syndrome: A 5-year follow-up study. *American Journal of Psychiatry* 158(12): 2038–2042.

Dowsett EG, Ramsay AM, McCartney RA, et al. (1990) Myalgic encephalomyelitis – a persistent enteroviral infection? *Postgraduate Medical Journal* 66(777): 526–530.

Dunning DL, Griffiths K, Kuyken W, et al. (2018) Research review: The effects of mindfulness-based interventions on cognition and mental health in children and adolescents – a meta-analysis of randomized controlled trials. *Journal of Child Psychology and Psychiatry* 60: 244–258.

Dutch Health Council (2018) ME/CVS Aan: de Voorzitter van de Tweede Kamer der Staten-Generaal Nr. 2018, Den Haag 19 maart 2018 *Gezondheidsraad*, 2018. Available at: https://huisartsvink.files.wordpress.com/2019/03/gr-kernadvies_me_cvs.pdf (accessed 27 March 2019)

Edwards J (2017) PACE team response shows a disregard for the principles of science. *Journal of Health Psychology* 22: 1155–1158.

Falk Hvidberg M, Brinth LS, Olesen AV, et al. (2015) The health-related quality of life for patients with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). *PLoS ONE* 10: e0132421.

Flo E and Chalder T (2014) Prevalence and predictors of recovery from chronic fatigue syndrome in a routine clinical practice. *Behaviour Research and Therapy* 63: 1–8.

Franklin JD, Atkinson G, Atkinson JM, et al. (2018) Peak oxygen uptake in chronic fatigue syndrome/myalgic encephalomyelitis: A meta-analysis. *International Journal of Sports Medicine* 40: 77–87.

Friedberg F and Krupp LB (1994) A comparison of cognitive behavioral treatment for chronic fatigue syndrome and primary depression. *Clinical Infectious Diseases* 18: S105–S110.

Friedberg F, Dechene L, McKenzie MJ, et al. (2000) Symptom patterns in long-duration chronic fatigue syndrome. *Journal of Psychosomatic Research* 48(1): 59–68.

Fukuda K, Straus SE, Hickie I, et al. (1994) The chronic fatigue syndrome: A comprehensive approach to its definition and study. *Annals of Internal Medicine* 121: 953–959.

Geraghty KJ (2017a) 'PACE-Gate': When clinical trial evidence meets open data access. *Journal of Health Psychology* 22(9): 1106–1112.

Geraghty KJ, Hann M and Kurtev S (2017b) Myalgic encephalomyelitis/chronic fatigue syndrome patients' reports of symptom changes following cognitive behavioural therapy, graded exercise therapy and pacing treatments: Analysis of a primary survey compared with secondary surveys. *Journal of Health Psychology*. Epub ahead of print 1 August. DOI: 10.1177/1359105317726152.

Ghatineh S and Vink M (2017) FITNET's internet-based cognitive behavioural therapy is ineffective and may impede natural recovery in adolescents with Myalgic Encephalomyelitis/ Chronic Fatigue Syndrome. A review. *Behavioral Science* 7(3): E52.

Green CR, Cowan P, Elk R', et al. (2014a) Draft executive summary: National Institutes of Health – Pathways to Prevention Workshop: Advancing the Research on Myalgic Encephalomyelitis/ Chronic Fatigue Syndrome. Available at: https://huisartsvink.files.wordpress.com/2018/08/green-draftreport-odp-mecfs-1.pdf (accessed 28 March 2019).

Green CR, Cowan P, Elk R', et al. (2014b) Final report: National Institutes Of Health – Pathways to Prevention Workshop: Advancing the Research on Myalgic Encephalomyelitis/ Chronic Fatigue Syndrome, executive summary. Available at: https://huisartsvink.files.wordpress.com/2018/08/green-odp-finalreport-p2p-mecfs.pdf (accessed 27 March 2019).

Haywood KL, Staniszewska S and Chapman S (2011) Quality and acceptability of patient-reported outcome measures used in chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ ME): A systematic review. *Quality of Life Research* 21(1): 35–52.

Heneghan C, Goldacre B and Mahtani KR (2017) Why clinical trial outcomes fail to translate into benefits for patients. *Trials* 18(1): 122.

Holmes GP, Kaplan JE, Gantz NM, et al. (1988) Chronic fatigue syndrome: A working case definition. *Annals of Internal Medicine* 108: 387–389.

Hróbjartsson A, Emanuelsson F, Thomsen AS, et al. (2014) Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *International Journal of Epidemiology* 43(4): 1272–1283.

Huibers MJ, Beurskens AJ, Van Schayck CP, et al. (2004) Efficacy of cognitive behavioural therapy by general practitioners for unexplained fatigue among employees: Randomised controlled trial. *British Journal of Psychiatry* 184: 240–246.

IOM (2015) Institute of Medicine; Committee on the Diagnostic Criteria for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome; Board on the Health of Select Populations. Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness. Washington, DC: National Academies Press, 2015.

Jason L, Torres-Harding S, Friedberg F, et al. (2007) Non-pharmacologic interventions for CFS: A randomized trial. *Journal of Clinical Psychology in Medical Settings* 172: 485–490.

Gibson-Saxty J (2002) Group therapy for chronic fatigue syndrome. National Research Register (this unpublished study could not be traced on the Internet).

Jones C, Hacker D, Xia J, et al. (2018) Cognitive behavioural therapy plus standard care versus standard care for people with schizophrenia. *Cochrane Database of Systematic Reviews, Issue* 12: CD007964.

Kennedy S (2002) Full remission: A return to normal functioning. *Journal of Psychiatry & Neuroscience* 27: 233–234.

Kindlon T (2011) Reporting of harms associated with graded exercise therapy and cognitive behavioural therapy in myalgic encephalomyelitis/chronic fatigue syndrome. *Bull IACFS ME* 19(2): 59–111. Available at: https://huisartsvink.files.wordpress.com/2018/08/kindlon-reporting-of-harms-associated-with-get-and-cbt-in-me-cfs.pdf (accessed 15 November 2018).

King MB (1999) Randomised controlled trial of counselling or CBT for patients with chronic fatigue in general practitioners. National Research Register.

Kirke KD (2017) PACE investigators' response is misleading regarding patient survey results. *Journal of Health Psychology* 22(9): 1168–1176.

Knoop H, Bleijenberg G, Gielissen MF, et al. (2007a) Is a full recovery possible after cognitive behavioural therapy for chronic fatigue syndrome? *Psychother Psychosom* 76(3): 171–176.

Knoop H, Prins JB, Stulemeijer M, et al. (2007b) The effect of cognitive behaviour therapy for chronic fatigue syndrome on self-reported cognitive impairments and neuropsychological test performance. *Journal of Neurology, Neurosurgery, and Psychiatry* 78(4): 434–436.

Komaroff AL, Fagioli LR, Doolittle TH, et al. (1996) Health status in patients with chronic fatigue syndrome and in general population and disease comparison groups. *The American Journal of Medicine* 101: 281–290.

Larun L, Brurberg KG, Odgaard-Jensen J, et al. (2017) Exercise therapy for chronic fatigue syndrome. *Cochrane Database of Systematic Reviews Issue* 4: CD003200.

Laws KR (2017) Distress signals: Does cognitive behavioural therapy reduce or increase distress in chronic fatigue syndrome myalgic encephalomyelitis? *Journal of Health Psychology* 22(9): 1177–1180.

Lilienfeld SO, Ritschel LA, Lynn SJ, et al. (2014) Why ineffective psychotherapies appear to work: A taxonomy of causes of spurious therapeutic effectiveness. *Perspectives on Psychological Science* 9: 355–387.

Lloyd AR, Hickie I, Brockman A, et al. (1993) Immunologic and psychologic therapy for patients with chronic fatigue syndrome: A double-blind, placebo-controlled trial. *American Journal of Medicine* 94: 197–203.

Lubet S (2017) Investigator bias and the PACE trial. *Journal of Health Psychology* 22(9): 1123–1127.

Luborsky L, Diguer L, Seligman DA, et al. (1999) The researcher's own therapy allegiances: A wild card in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice* 6(1): 95–106.

Luborsky L, Rosenthal R, Diguer L, et al. (2002) The dodo bird verdict is alive and well – Mostly. *Clinical Psychology: Science and Practice* 9(1): 2–12.

Lundh A, Sismondo S and Lexchin J (2012) Industry sponsorship and research outcome. *Cochrane Database of Systematic Reviews* 12: MR000033.

McCrone P, Sharpe M, Chalder T, et al. (2012) Adaptive pacing, cognitive behaviour therapy, graded exercise, and specialist medical care for chronic fatigue syndrome: A cost-effectiveness analysis. *PLoS ONE* 7: e40808.

Marks DF (2017) Special issue on the PACE Trial. *Journal of Health Psychology* 22(9): 1103–1105.

Meyer K, Schwaibolda M, Westbrook S, et al. (1997) Effects of exercise training and activity restriction on 6-minute walking test performance in patients with chronic heart failure. *American Heart Journal* 133: 447–453.

Mohr DC, Spring B, Freedland KE, et al. (2009) The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics* 78(5): 275–284.

Morriss RK, Wearden AJ and Mullis R (1998) Exploring the validity of the Chalder Fatigue scale in chronic fatigue syndrome. *Journal of Psychosomatic Research* 45(5): 411–417.

Moss-Morris R, Sharon C, Tobin R, et al. (2005) A randomized controlled graded exercise trial for chronic fatigue syndrome: Outcomes and mechanisms of change. *Journal of Health Psychology* 10(2): 245–259.

Munder T, Brütsch O and Leonhart R (2013) Researcher allegiance in psychotherapy outcome research: An overview of reviews. *Clinical Psychology Review* 33(4): 501–511.

Munder T, Flückiger C, Gerger H, et al. (2012) Is the allegiance effect an epiphenomenon of true efficacy differences between treatments? A meta-analysis. *Journal of Counselling Psychology* 59(4): 631–637.

NIH Project Reporter (2008) Research Portfolio Online Reporting Tools (RePORT) Project Number: 5R21NR010539-02 Former Number: 1R21AR054506-01A1 Contact PI/Project Leader: Krystal, A.D., Duke University, Title: Behavioral Insomnia Treatment in Chronic Fatigue. https://projectreporter.nih.gov/project_info_description.cfm?aid=7473928&icde=41688725&ddparam=&ddvalue=&ddsub=&cr=2&csb=default&cs=ASC&pball= (accessed 15 November 2018).

Núñez M, Fernández-Solà J, Nuñez E, et al. (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. *Clinical Rheumatology* 30: 381–389.

O'Dowd H, Gladwell P, Rogers C, et al. (2006) Cognitive behavioural therapy in chronic fatigue syndrome: A randomised controlled trial of an outpatient group programme. *Health Technology Assessment* 10(37): 1–121.

O'Dowd H (2000) Cognitive behavioural therapy in chronic fatigue syndrome (CFS): A randomised controlled trial of an outpatient group programme. *Current Controlled Trials* 10: 1–21.

O'Leary D (2018) Bodily distress syndrome: Concerns about scientific credibility in research and implementation. *Journal of Biological Physics and Chemistry* 18(2018): 67–77.

Olesen AV, Oddershede L and Petersen KD (2016) Health-related quality of life in Denmark on a relative scale: Mini-catalogue of mean EQ-5D-3L index scores for 17 common chronic conditions. *Nordic Journal of Health Economics* 4(2): 44–56.

Pendergrast T, Brown A, Sunnquist M, et al. (2016) Housebound versus nonhousebound patients with myalgic encephalomyelitis and chronic fatigue syndrome. *Chronic Illness* 12: 292–307.

Price JR, Mitchell E, Tidy E, et al. (2008) Cognitive behaviour therapy for chronic fatigue syndrome in adults (Review). *Cochrane Database of Systematic Reviews* 4: CD001027

Prins JB, Bleijenberg G, Bazelmans E, et al. (2001) Cognitive behaviour therapy for chronic fatigue syndrome: A multicentre randomised controlled trial. *The* Lancet 357: 841–847.

Reeves WC, Lloyd A, Vernon SD, et al. (2003) Identification of ambiguities in the 1994 chronic fatigue research case definition and recommendations for resolution. *BMC Health Services Research* 3: 25.

Ridsdale L, Darbishire L and Seed PT (2004) Is graded exercise better than cognitive behaviour therapy for fatigue? A UK randomized trial in primary care. *Psychological Medicine* 34: 37–49.

Ridsdale L, Godfrey E, Chalder T, et al. (2001) Chronic fatigue in general practice: Is counselling as good as cognitive behaviour therapy? *British Journal of General Practice* 51(462): 19–24.

Roberts D (2018) Chronic fatigue syndrome and quality of life. *Patient Related Outcome Measures* 9: 253–262.

Rongen-van Dartel SA, Repping-Wuts H, van Hoogmoed D, et al. (2014) Relationship between objectively assessed physical activity and fatigue in patients with rheumatoid arthritis: Inverse correlation of activity and fatigue. *Arthritis Care Res (Hoboken)* 66(6): 852–860.

Russell V, Gaston AM, Lewin RJP, et al. (2001) Group rehabilitation for adult chronic fatigue syndrome. Unpublished Article.

Scheeres K, Knoop H, van der Meer J, et al. (2009) Clinical assessment of the physical activity pattern of chronic fatigue syndrome patients: A validation of three methods. *Health Qual Life Outcomes* 7: 29.

Sciurba F, Criner GJ, Lee SM, et al. (2003) National Emphysema Treatment Trial Research Group. Six-minute walk distance in chronic obstructive pulmonary disease: Reproducibility and effect of walking course layout and length. *American Journal of Respiratory and Critical Care Medicine* 167(11): 1522–1527.

Sharpe M (1993) Non-pharmacological approaches to treatment. *Chronic Fatigue Syndrome* 173: 298–308.

Sharpe M, Goldsmith KA and Johnson AL (2015) Rehabilitative treatments for chronic fatigue syndrome: Long-term follow-up from the PACE trial. *The Lancet* 2: 1067–1074.

Sharpe M, Hawton K, Simkin S, et al. (1996) Cognitive behaviour therapy for the chronic fatigue syndrome: A randomized controlled trial. *British Medical Journal* 312: 22–26.

Sharpe MC, Archard LC, Banatvala JE, et al. (1991) A report – chronic fatigue syndrome: Guidelines for research. *Journal of the Royal Society of Medicine* 84(2): 118–121.

Smith MEB, Nelson HD, Haney E, et al. (2016) Diagnosis and treatment of myalgic encephalomyelitis/chronic fatigue syndrome. Evidence Reports/Technology Assessments, No. 219. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2012-00014-I.) AHRQ Publication No. 15-E001-EF. Rockville, MD: Agency

for Healthcare Research and Quality; December 2014. Addendum July 2016. Available at: https://huisartsvink.files.wordpress.com/2018/08/ahrq-smith-et-al-2016-chronic-fatigue_research.pdf (accessed 4 November 2018).

Stevens D, Elpern E, Sharma K, et al. (1999a) Comparison of hallway and treadmill six-minute walk tests. *Am J Respir Crit Care Med* 160: 1540–1543.

Stevens M, Gevirtz R, Verity L, et al. (1999b) Chronic fatigue syndrome: A chronobiologically oriented controlled treatment outcome study (this unpublished study cannot be traced on the Internet).

Stordeur S, Thiry N and Eyssen M (2008) Chronisch Vermoeidheidssyndroom: Diagnose, behandeling en zorgorganisatie [Fatigue Syndrome: Diagnosis, treatment and organisation of care]. Technical report 88A (in Dutch). Brussels: Belgian Healthcare Knowledge Center (KCE). Available at: https://kce.fgov.be/sites/default/ files/page_documents /d20081027358.pdf (accessed 15 November 2018)

Strang JM (2002) Treatment of chronic fatigue syndrome: A cognitive-behavioral approach to enhance personal mastery. *Dissertation Abstracts International Section A: Humanities and Social Sciences,* 63(5-A), 1728. Available at: http://psycnet.apa.org/record/2002-95021-130 (accessed 15 November 2018)

Stulemeijer M, De Jong LW, Fiselier TJ, et al. (2005) Cognitive behaviour therapy for adolescents with chronic fatigue syndrome: Randomised controlled trial. *BMJ* 330(7481): 14. Erratum in Science 9: 355, BMJ 330(7495): 820.

Surawy C, Roberts J and Silver A (2005) The effect of mindfulnes training on mood and measures of fatigue, activity, and quality of life in patients with chronic fatigue syndrome on a hospital waiting list: A series of exploratory studies. *Behavioural and Cognitive Psychotherapy* 33: 103–109.

Tolin DF (2012) Is cognitive–behavioral therapy more effective than other therapies? A meta-analytic review. *Clinical Psychology Review* 30: 710

Turner EH, Matthews AM, Linardatos E, et al. (2008) Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 358(3): 252–260

Twisk FNM and Corsius LAMM (2017) 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. *Gen Med Open* 1(3): 1–13.

Van den Berg-Emons RJ, Ortye L, Buffart AA, et al. (2011) Validation of the Physical Activity Scale for individuals with physical disabilities. *Archives of Physical Medicine and Rehabilitation* 92: 923–928.

Van Dessel N, Den Boeft M, van der Wouden JC, et al. (2014) Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults. *Cochrane Database of Systematic Reviews* 10: CD011142.

Van Kessel K, Moss-Morris R, Willoughby L, et al. (2008) A randomized controlled trial of cognitive behavior therapy for multiple sclerosis fatigue. *Psychosomatic Medicine* 70: 205–213.

Vink M (2016) The PACE trial invalidates the use of cognitive behavioral and graded exercise therapy in myalgic encephalomyelitis/chronic fatigue syndrome: A review. *Neurology and Neurobiology Journal* 2(3): 112–130.

Vink M (2017a) Assessment of individual PACE trial data: In myalgic encephalomyelitis/chronic fatigue syndrome, cognitive behavioral and graded exercise therapy are ineffective, do not lead to actual recovery and negative outcomes may be higher than reported. *Neurology and Neurobiology Journal* 3(1): 02813.

Vink M (2017b) PACE trial authors continue to ignore their own null effect. *Journal of Health Psychology* 27: 1–7.

Vink M and Vink-Niese A (2018a) Multidisciplinary rehabilitation treatment is not effective for myalgic encephalomyelitis/chronic fatigue syndrome: A review of the FatiGo trial. *Health Psychology* 12: 1–8.

Vink M and Vink-Niese A (2018b) Graded exercise therapy for myalgic encephalomyelitis/chronic fatigue syndrome is not effective and unsafe. Re-analysis of a Cochrane review. *Health Psychology Open* 8: 5.

Vissers H (2008) The efficacy and predicting variables of a multidisciplinary disability resolution (MDR) program for CFS patients receiving long term disability benefits from income protection insurers. Controlled-trials.com (this unpublished study cannot be traced on the Internet).

Wearden AJ (2001) Protocol for the (FINE) trial: A randomised controlled trial of nurse facilitated self-help treatment for patients in primary care with chronic fatigue syndrome ISRCTN74156610 https://huisartsvink.files.wordpress.com/2019/03/wearden-2001isrctn.pdf (accessed 27 March 2019).

Wearden AJ and Emsley R (2013) Mediators of the effects on fatigue of pragmatic rehabilitation for chronic fatigue syndrome. *Journal of Consulting and Clinical Psychology* 81: 831–838.

Wearden AJ, Dowrick C, Chew-Graham C, et al. (2010a) Fatigue scale. *BMJ Rapid Response*, 27 May. Available at: http://www.bmj.com/rapid-response/2011/11/02/fatigue-scale-0 (accessed 15 November 2018)

Wearden AJ, Dowrick C, Chew-Graham C, et al. (2010b) Fatigue Intervention by Nurses Evaluation (FINE) trial writing group and the FINE trial group. Nurse led, home based self-help treatment for patients in primary care with chronic fatigue syndrome: Randomised controlled trial. *BMJ* 340: 1777.

Wearden AJ, Riste L, Dowrick C, et al. (2006) Fatigue Intervention by Nurses Evaluation – The FINE Trial. A randomised controlled trial of nurse led self-help treatment for patients in primary care with chronic fatigue syndrome: Study protocol. *BMC Medicine* 4: 9.

Wechsler ME, Kelley JM, Boyd IO, et al. (2011) Active albuterol or placebo, sham acupuncture, or no intervention in asthma. *New England Journal of Medicine* 365(2): 119–126.

White (2005) The corresponding reference is White (2007).

White PD (2003) RCT of CBT, graded exercise, and pacing versus usual medical care for the chronic fatigue syndrome ISRCTN54285094, 22 May. Available at: https://huisartsvink.files.wordpress.com/2018/09/pace-isrctn54285094.pdf (accessed 15 November 2018).

White PD, Goldsmith K, Johnson AL, et al. (2013) Recovery from chronic fatigue syndrome after treatments given in the PACE trial. *Psychological Medicine* 43: 2227–2235.

White PD, Goldsmith KA, Johnson AL, et 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. *The Lancet* 377: 823–836.

White PD, Sharpe MC, Chalder T, et al. (2007) Protocol for the PACE trial: A randomised controlled trial of adaptive pacing, cognitive behaviour therapy and graded exercise as supplements to standardised specialist medical care versus standardised specialist medical care alone for patients with the chronic fatigue syndrome/myalgic encephalomyelitis or encephalopathy. *BMC Neurology* 7: 6.

Whitehead L and Campion P (2002) Can general practitioners manage chronic fatigue syndrome. *Journal of Chronic Fatigue Syndrome* 10(1): 55–64.

Whiting P, Bagnall AM, Sowden AJ, et al. (2001) Interventions for the Treatment and Management of Chronic Fatigue Syndrome – A Systematic Review. *JAMA* 286: 1360–1368.

World Health Organization [WHO] (2010) International Statistical Classification of Diseases and Health Related Problems (the) ICD-10, Volume 1: Tabular List, 2nd ed.; WHO: Geneva, Switzerland, 10th revision Volume 3 Alphabetical index Fifth edition.

Wiborg JF, Knoop H, Stulemeijer M, et al. (2010) How does cognitive behaviour therapy reduce fatigue in patients with chronic fatigue syndrome? The role of physical activity. *Psychological Medicine* 40: 1281–1287.

Wiborg JF, van Bussel J, van Dijk A, et al. (2015) Randomised controlled trial of cognitive behaviour therapy delivered in groups of patients with chronic fatigue syndrome. *Psychother Psychosom* 84(6): 368–376.

Wilshire C (2017) The problem of bias in behavioural intervention studies: Lessons from the PACE trial. *Journal of Health Psychology* 22(9): 1128–1133.

Wilshire CE, Kindlon T, Courtney R, et al. (2018b) Rethinking the treatment of chronic fatigue syndrome – a reanalysis and evaluation of findings from a recent major trial of graded exercise and CBT. *BMC Psychology* 22(1): 66.

Wilshire CE, McPhee G, the Science for MECFQ, et al. (2018a) Submission to the public review on common data elements for ME/CFS: Problems with the Chalder Fatigue Questionnaire. *Science for ME January* 22. Available at: https://huisartsvink.files.wordpress.com/2018/08/wilshire-mcphee-cfq-cde-critique-for-s4me-final.pdf (accessed 15 November 2018).

Appendix 1

Analysis of trials that were still in progress

According to the reviewers, there were six trials still in progress (Bleijenberg, 2008; Carney and Jones-Alexander 2008; Gibson-Saxty, 2002; Vissers, 2008; Wearden et al., 2006; White, 2005), which will help to strengthen the evidence base on CBT interventions for CFS. They wrote that in 2008. These trials should have been published by now and we will analyse them here.

Bleijenberg (2008) (study protocol), published as Wiborg et al. (2015)

This trial was not properly controlled: 14 sessions (2 hours of CBT over a 6-month period; split into groups of four and eight patients) and 0 (waiting list control group). The authors acknowledged that this could artificially inflate the treatment effect; 37.3 per cent (124/328) of patients who were asked to engage in group therapy refused that so that only those who were willing and motivated to have group therapy were enrolled into the trial, 19 per cent (32/168; CBT) and 11.8 per cent (8/68; waiting list) of participants dropped out. Missing data were imputed using mean proportions of improvement based on the outcome scores of similar patients with a second assessment. This might have artificially inflated the results, because participants who do not respond to treatment or are negatively affected by it, are more likely to drop out or be lost to follow-up (Lilienfeld et al., 2014).

Psychological distress improved by 16.7 per cent (27/162, CBT) and 10.5 per cent (18/171, waiting list), suggesting that the subjective improvement in the treatment group might be a reflection of improvement in psychological well-being.

The study used a post hoc definition of clinically significant improvement and recovery using 'rigorous criteria for normal functioning'. They concluded that more people had recovered after CBT (15.4%) than in the waiting list group (1.5%). However, in a well-designed trial, such a definition should be defined before it starts, to avoid that the results can influence it. Also they did not use rigorous criteria. One was deemed to have recovered with the following scores: 80 or more (physical functioning), less than 27 (CIS-fatigue) and less than 203 (SIP overall impairment), even though healthy people, with a similar mean age, have scores of 93.1 (PF), 17.3 (CIS-fatigue) and 65.5 (SIP) according to Knoop et al. (2007a), which was used by the authors.

In view of the above-mentioned problems, and not using objective outcomes, one cannot safely conclude that group CBT is effective and leads to recovery in one in eight patients.

Carney and Jones-Alexander (2008)

This project received two grants, totalling $418,335 (NIH Project Reporter, 2008), but no results have been published yet.

Vissers (2008)

This study has not been published yet.

Wearden et al. (2006) (study protocol), published as Wearden et al. (2010b)

Wearden et al. (2010b) is often referred to as the FINE trial (**F**atigue **I**ntervention by **N**urses **E**valuation). This trial was not properly controlled. The average number of sessions in the trial was as follows: 9.63 (pragmatic rehabilitation group (PR)), 9.5 (supportive listening group (SL)) and 0 (no treatment group) (treatment as usual by their own GP when needed). Number of GP consultations: 2 (PR), 3 (SL) and 3 (no treatment).

The entry criteria were changed from the Fukuda to the even wider Oxford criteria, 8 months into a 4-year lasting trial. No reason was given (Wearden, 2001). According to the FINE trial protocol (Wearden et al., 2006), primary outcomes were to be self-reported physical functioning and fatigue at 1 year. Yet in the 2010 paper, this was changed to 20 weeks (end of treatment) and 70 weeks from recruitment; 20.3 per cent (60/296) of the participants suffered from anxiety and 17.9 per cent (53/296) from depression.

The only objective secondary outcome measure (step test) was omitted from the 2010 paper even though not publishing results jeopardises the validity of a study (Heneghan et al., 2017). When the results were published 3 years later (Wearden and Emsley, 2013), there were no differences between pragmatic rehabilitation and no treatment. The fatigue scores were changed from bimodal (0–11) to Likert (0–33) in a Rapid Response in the *BMJ* (Wearden et al., 2010a) and in Wearden and Emsley, 2013. This change was made despite the fact that two of the authors (including the Principal Investigator) concluded in a paper, devoted to analysing the use of the Chalder Fatigue scale in CFS (Morriss et al., 1998), that near-maximal scoring on six physical fatigue scale items from the total of 14 items (five if the 11 item scale is used), supports using the 2-point bimodal, rather than the 4-point Likert scoring. Re-scored there was now a clinically modest, but statistically significant effect of PR compared with no treatment at both outcome points. However, altering measures in this way after the trial to find a small effect suggests a form of p-hacking.

The entry criteria, outcome switching and null objective improvement in this trial mean that it is unsafe to claim any effect for the interventions.

The PACE trial

White (2005), the corresponding reference is White et al. (2007) (study protocol), published as White et al. (2011). This trial is the largest CBT and graded exercise therapy (GET) study conducted so far. It included 640 patients, compared to 1008 in the Cochrane review. This study used the Oxford criteria, 47 per cent had a comorbid depression or anxiety disorder and 80 per cent of the screened patients were not selected. The control group did not have the same number of contact hours: 16 sessions (CBT), 17 sessions GET (both including 3 sessions of SMC), yet the SMC group only had 5. This imbalance creates serious biases towards finding a positive effect for the intervention, regardless of whether it's effective or not (Lilienfeld et al., 2014).

A null effect at long-term follow-up was spun as positive. Outcomes with SMC alone or adaptive pacing therapy (APT) improved from the 1 year outcome and were similar to CBT and GET at long-term follow-up, but it was claimed the data should be interpreted in the context of additional therapies having been given after the 1-year trial final assessment (Sharpe et al., 2015). However, the Supplementary appendix long-term follow-up shows that the majority of participants did not have any additional CBT (76%) or GET (83%) after the trial. It also shows that patients in all four groups, who did not receive additional treatment subsequent to trial completion, exhibited lower fatigue and higher physical functioning scores relative to those of patients who received additional treatment (Vink, 2016).

Baseline figures were used for one objective test, an actometer, a reliable measure of activity to assess improvement objectively (Scheeres et al., 2009), but were not recorded at the end of the trial. The reason given was that it would be too great a burden for patients (Vink, 2016), even though they had consented to use it, they had completed moderately effective treatment (White et al., 2011) and 22 per cent of those in the CBT and GET groups had recovered according to the investigators (White et al., 2013).

An extensive number of endpoint changes were made (Sharpe et al., 2015; Vink, 2016; White et al., 2011; Wilshire et al., 2018b). The timing of the changes to the primary outcomes – several months after trial completion – was highly problematic (Wilshire et al., 2018b). As a result, there was suddenly an overlap in entry and recovery criteria: 13.3 per cent of participants were already recovered according to one (12.8%) or two (0.5%) of the recovery criteria at trial entry (Vink, 2017a). That is before receiving any treatment and without a change in their medical situation.

These changes affected both the physical function scores (PF) and the fatigue scores. The minimum PF required to qualify as recovered was reduced from 85 to 60 (White et al., 2011). The maximum score for trial entry was increased from 60 to 65 (0–100; higher scores indicating better functioning). Even though according to the PACE trial's recovery article, a score of 65 or less represents 'abnormal levels of physical function' (White et al., 2013) and severe disability according to the literature (Stulemeijer et al., 2005). Participants with a score of 60 to 65 (inclusive) were thus considered ill enough to participate and to have an abnormal level of physical functioning, yet were also recovered and severely disabled. Three participants (0.45%) saw their physical functioning score go down from 65 to 60, reflecting deterioration, and three others (0.45%) had unchanged physical functioning scores, but all (0.9%) were still classed as recovered, according to the physical functioning recovery criterion (Vink, 2017b).

Something similar happened to the fatigue scores. When PACE was registered with the ISRCTN on 22 May 2003, participants needed a Chalder Fatigue Questionnaire (CFQ) score of four or more to be classed as ill enough to take part (White, 2003). The CFQ entry criterion was changed to six or more before the trial started and then during a non-blinded trial switched from bimodal to Likert, 18 or more to qualify. To be classed as recovered, a bimodal score of $\leqslant 3$ out of 11, which represented a screening threshold for abnormal fatigue, was changed to a Likert score of 18 or

less (0–33) (White et al., 2013). Consequently, with a Likert score of 18, one was simultaneously classed as disabled and recovered. These endpoint changes increased recovery rates of CBT and GET fourfold. Had the PACE trial stuck to the protocol defined endpoints, then there would have been no statistically significant difference in recovery rates between the four treatment groups (Wilshire et al., 2018b).

The net improvement of the quality of life scores (EQ-5D) after CBT at 52 weeks over APT was 1.8 per cent (0.09/0.63 – 0.06/0.48) (McCrone et al., 2012). A study by Olesen et al. (2016) of 20,220 adult patients found a mean quality of life score of 0.84 for the total population and 0.93 for people without a chronic condition. Yet the quality of life at 52 weeks in the CBT group (0.63) (McCrone et al., 2012) was similar to the score for cerebral thrombosis (0.62) and still worse than in rheumatoid arthritis and angina (0.65), AMI (acute myocardial infarction) (0.66) (Olesen et al., 2016), MS (0.67), lung cancer and people with four or more chronic health conditions (0.69), stroke (0.71) or ischemic heart disease (0.72) (higher scores indicating a better quality of life) (Falk Hvidberg et al., 2015). Also, there was no statistically significant difference in the improvement in CFS symptom count between CBT and APT ($p = 0.0986$) at 52 weeks.

These flaws in the trial render unsafe any conclusion that CBT is effective.

Review of the objective outcomes

Two of the trials that were still in progress when the Cochrane review was published (Wearden et al., 2010b; White et al., 2011) used objective outcomes too. They form an important part of the evidence base as combined they included 936 (640 + 296) participants. The scores for the only objective outcome used in Wearden et al. (2010b), the step test, showed no differences between the pragmatic rehabilitation and no treatment (GP treatment as usual) groups on any of the step test measures at 20 or 70 weeks (Wearden and Emsley, 2013).

In White et al. (2011), the step test did not show any objective improvements, therefore fitness did not improve. This is matched by the net improvement of the quality of life scores after CBT over APT of only 1.8 per cent. The number of patients who were unable to work and who were receiving benefits increased and the number of patients receiving income protection in the CBT group actually doubled (McCrone et al., 2012). In addition, there was no statistically significant difference in the improvement in CFS symptom count between CBT and APT ($p = 0.0986$) at 52 weeks. There was no statistical significant difference in the 6-minute walking test outcome between CBT and SMC ($p = 0.87$) and CBT and APT ($p = 0.65$) after exercising for 24 weeks, at 52 weeks. According to these results − 354 m after CBT – patients would still be ill enough to be put on the waiting list for a lung transplant ($\leqslant 400$ m) (Vink, 2016). No one in the trial achieved actual recovery, where symptoms are eliminated and patients return to pre-morbid levels of functioning (Kennedy, 2002), which is the general public's understanding of the meaning of recovery (Vink, 2017a). The PACE trial protocol defined improvement as an increase of 50 per cent. According to the 6-minute walk test results, the only objective individual results that were released, this benchmark was matched by 3.7 per cent in the CBT group, but also by 5 per cent in the SMC group, implying a negative effect of CBT of 1.3 per cent (participants in all treatment groups also received SMC) (Vink, 2017a).

Quality of life

In White et al. (2011), participants improved 1.8 per cent more after CBT then after APT at 52 weeks (McCrone et al., 2012).

Employment status

In White et al. (2011), the number of patients who were unable to work and who were receiving benefits, due to illness or disability, increased and the number of patients receiving income protection in the CBT group doubled (McCrone et al., 2012).

Appendix 2

Support for their recovery claims

Two studies, Knoop et al. (2007a) and Flo and Chalder (2014), were used by White et al. (2011) as support for their recovery claims. We will therefore analyse these two studies too.

Knoop et al. (2007a)

This non-randomised cohort trial had no control group and it did not use objective outcomes. The authors concluded that 23 per cent had 'fully recovered' after CBT. They acknowledged that in the absence of a control group, it is difficult to attribute this to treatment with certainty. The recovery definition included having a maximum CIS-fatigue score of 27, a minimum physical function score of 80, a maximum SIP functional disability score of 203 and a minimum health perception score of 65, which is the score of people aged 65–74 (Twisk and Corsius, 2017). The mean age in this trial is 37.0. However, according to the authors, the scores for healthy people of a similar age are the following: 17.3 (CIS-fatigue), 93.1 (PF), 80 (health perception) and 65.5 (SIP).

In all, 5.2 per cent (5/96) had no PEM. According to the authors, some suggest that this is the main characteristic feature of CFS. It is unclear why these people were not excluded from the study. Drop-out rate was 11 per cent.

One cannot come to any meaningful conclusions about the efficacy of CBT in a non-blinded non-randomised trial without a control group that uses a broad definition of recovery and does not use objective outcomes.

Flo and Chalder (2014)

This non-randomised cohort trial did not use a control group or objective outcomes; 28.9 per cent already fulfilled the physical functioning recovery criterion at trial entry. It used the English NICE criteria (having fatigue for the last 4 months) – which are even wider than the Oxford criteria − 72.7 per cent fulfilled the Oxford and 52.6 per cent the Fukuda criteria.

The recovery criteria were very wide; for example, one was classified as recovered with a physical functioning score of 65 or more (0–100; higher score means better physical function). The problems with choosing a score that low have been discussed earlier; 28.9 per cent had already achieved this at trial entry, 12.3 per cent already had a score of 83 or more at trial entry, just like 16.1 per cent already fulfilled the Chalder fatigue questionnaire recovery score (<18; range 0–33) at trial entry. Drop-out rate was 27.8 per cent. The authors used Knoop et al. (2007a; 23%) and White et al. (2011; 22%) as support for their claim that 18.3 per cent had 'fully recovered' after CBT. No objective outcomes or a control group were used in this non-blinded non-randomised cohort trial. One cannot sustain the recovery claim in view of all these flaws.

Appendix 3

Núñez et al. (2011)

This study was excluded from a recent Cochrane exercise review, as found by a reanalysis (Vink and Vink-Niese, 2018), because exercise therapy was a minor part of the intervention and it did not measure fatigue, viewed as primary outcome in the review.

The trial compared multidisciplinary treatment combining CBT, GET and pharmacological treatment with usual treatment, with 1-year follow-up after the end of treatment. It concluded that at 12 months, the interventions did not improve health-related quality of life scores, and led to worse physical function and bodily pain scores. Núñez et al. found that the combination of CBT and GET is ineffective and not evidence-based and may in fact be harmful.

Table 2. Summary of the findings of the trials published since the Cochrane review.

Studies	Treatment	No. of participants	Selection criteria	Control group	Psychiatric comorbidity	Objective outcomes/ work	Quality of life	Missing data/ dropouts
Carney and Jones-Alexander, 2008	Unpublished study							
Flo and Chalder, 2014	CBT	200 (non-randomised cohort trial; 28.9% and 16.1% fulfilled physical functioning and fatigue recovery score, respectively, at trial entry)	NICE	No control group	No information provided	Not used	Not used	27.8% (dropouts)
Gibson-Saxty, 2002	Unpublished study							
Knoop et al., 2007a	CBT	96 (non-randomised cohort trial; post hoc definition of recovery; 5% no PEM)	Fukuda	No control group	No information provided	Not used	Not used	11% (dropouts)
Núñez et al., 2011	Multidisciplinary treatment consisting of group CBT, GET and conventional pharmacological symptomatic treatment	120	Fukuda	Usual CFS therapy including exercise counselling and conventional pharmacological symptomatic treatment	19% anxiety and 20% depression	Not used	No information	4% (treatment group) and 5% (control group) dropped out
Vissers, 2008	Unpublished study							
Wearden et al., 2010b	Pragmatic rehabilitation and supportive listening	296 (a number of endpoint changes were made; one after the trial had been published)	Oxford	No treatment	29.7% anxiety and/or depression	No improvement (step test)	Not used	13% of patients dropped out
White et al., 2011	CBT and GET	640 (extensive number of endpoint changes were made; consequently, 13.3% were already recovered according to 1 or 2 recovery criteria at trial entry)	Oxford	SMC and APT, not properly controlled	47% anxiety and/or depression	No improvement (6-minute walk test and step test); work status and sickness and disability status deteriorated	+1.8% compared to APT	11% CBT, 7% APT (dropouts), 24% CBT, 26% SMC (missing 6-minute walk test data)
Wiborg et al., 2015	Group CBT	204	Fukuda	Waiting list, not properly controlled	No information provided (trained therapists ruled out psychiatric comorbidity as potential explanation for the complaints)	Not used	Not used	19% CBT and 11.8% waiting list (missing data)

CBT: cognitive-behavioural therapy; CFS: chronic fatigue syndrome; GET: graded exercise therapy; NICE: National Institute for Health and Care Excellence; PEM: postexertional malaise; APT: adaptive pacing therapy.