Do more people recover from chronic fatigue syndrome with cognitive behaviour therapy or graded exercise therapy than with other treatments?

Sharpe M¹, Chalder T², Johnson AL³, Goldsmith KA⁴, White PD⁵

1 Psychological Medicine Research, University of Oxford Department of Psychiatry, Oxford, UK
2 Academic Department of Psychological Medicine, King’s College London, London, UK
3 MRC Clinical Trials Unit, London, UK
4 Biostatistics & Health Informatics Department, Division of Psychology and Systems Sciences, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
5 Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, UK

Correspondence to: Professor M Sharpe, Psychological Medicine Research, University of Oxford Department of Psychiatry, Oxford, UK, michael.sharpe@psych.ox.ac.uk
Abstract

BACKGROUND: Wilshire and colleagues suggest that we have overestimated the number of patients that recover from chronic fatigue syndrome (CFS) after receiving a course of either cognitive behaviour therapy (CBT) or graded exercise therapy (GET), as reported in a secondary analysis of outcome data from the PACE trial. We provide counter-arguments to this view.

PURPOSE: To provide an alternative view to that offered by Wilshire and colleagues.

METHODS: We review the two different analyses of PACE trial outcome data to estimate recovery, and compare these to other published trials and cohort studies.

RESULTS: While there is no gold standard for recovery from chronic fatigue syndrome, previous trials and cohort studies provide support for our analysis of recovery rates in the PACE trial. Our finding that 22% of participants recover from the current episode of CFS after either CBT or GET, compared to 8% after adaptive pacing therapy and 7% after specialist medical care alone, is consistent with the literature.

CONCLUSIONS: Our original conclusions that “...recovery from CFS is possible, and that CBT and GET are the therapies most likely to lead to recovery” is an important, reasonable, and hopeful message for both clinicians and patients.

Keywords: Chronic fatigue syndrome, myalgic encephalomyelitis, cognitive behaviour therapy, graded exercise therapy, recovery

Text – 1531 words
The PACE trial found that, when added to specialist medical care, both cognitive behaviour therapy (CBT) and graded exercise therapy (GET) are safe treatments for chronic fatigue syndrome (CFS), and are more effective than both adaptive pacing therapy (APT) added to SMC and SMC alone [1]. These findings emerge both from our planned detailed statistical analysis [1, 2], and from the analysis originally outlined in the published protocol [3-5]. Furthermore, an updated systematic review of all eight trials of exercise therapy for CFS, which included the PACE trial data, confirmed that exercise therapy is an effective and safe treatment for CFS, and that it is as effective as CBT [6].

So, why do Wilshire and colleagues cast doubt on these findings, and even suggest that clinicians and patients have been misled [7]? It is important to note that their argument focusses not on the primary trial outcomes, outlined above, but on one of a number of secondary analysis of the PACE trial data, which explored whether patients were more likely to recover from their current episode of CFS after a course of either CBT or GET, compared to APT and SMC [8]. Our analysis suggested that they were more likely to recover, whilst Wilshire and colleagues suggest they were not. The question we address here is which view is right, given that there is no generally agreed measure of recovery.

On first glance it may appear that we were wrong. In particular that got our chosen thresholds wrong, as in our analysis 13% of participants met the recovery criterion of being within the normal range (population mean, minus one standard deviation) for physical functioning when they entered the trial. This observation would support the argument that this threshold we chose was too lenient to be used to assess recovery. However, further scrutiny shows that this argument is misleading as being within normal range on this single measure was only one of the criteria we used to define recovery; to be regarded as recovered participants also had to meet four additional criteria [8]. These were: (1) being within the population normal range for fatigue; (2) being within the normal range for both fatigue and
physical functioning (only 3 out of 640 (<1%) patients satisfied this criterion at baseline); (3) no longer meeting the Oxford definition for CFS, which included no longer meeting trial severity eligibility for entry into the trial for either fatigue (binary score of 5 or fewer out of 11) or physical function (score of 70 or more out of 100) (no patient met this criterion at baseline); and (4) self-rating their overall health (not just CFS) as “very much better” or “much better” compared to trial entry (by definition no one met this criterion at baseline).

Therefore, even if it is argued that the threshold we used for physical function was set too low, no participant met our full criteria for recovery at baseline.

Our findings differ from those of Wilshire’s co-authors’ who report that only 7%, 4%, and 3% of participants recovered for CBT, GET, and SMC respectively [7, 9]. They don’t mention the figure for APT, which was 2% [9]. Why is there such discrepancy between these two analyses? The main reason is that the analysis adopted by Wilshire and colleagues used thresholds that we proposed in the outline analysis plan taken from our published protocol [3]. These were: fatigue binary score of < 4/11, SF36 physical function score of >80/100, not meeting the Oxford definition of having CFS, and participant rating of overall health as “very much better” [8]. We changed these thresholds for our detailed analysis plan because, after careful consideration and consultation, we concluded that they were simply too stringent to capture clinically meaningful recovery [8]. For instance, we decided that rating one’s overall health as “much better” or “very much better” was more consistent with recovery on this criterion than only “very much better”. Another more minor reason may be that Wilshire’s and colleagues imputed missing data by last observation carried forward, a method likely to reduce the observed rate of recovery [7, 9].

These differences in definition of recovery make it unsurprising that the two papers report different results. And as we wrote in our paper, “the main limitation of this analysis is the absence of a generally agreed measure of recovery” [8]. This caveat of course also applies to
the analysis done by Matthees and colleagues. Given that there is no ‘right answer’ to what is recovery, we might reasonably ask which of the two definitions of recovery is most clinically valid.

**What have others studies found regarding recovery?**

One way to address the question of the clinical validity of each definition of recovery is to review the recovery rates reported from naturalistic cohort studies. Cairns and Hotopf [10] systematically reviewed follow-up studies (the majority not involving any treatment) of patients with operationally defined CFS, and found a median (range) of only 7% (0–48%) recovered over time. This percentage is the same as the one we observed following SMC in the PACE trial, and provides some support for the validity of this finding.

We can also compare results from a previous trial of CBT. In a five year follow-up study of a trial comparing CBT to relaxation therapy, Deale and colleagues [11] found that 24% of 25 patients met operationalised criteria for recovery at 5 years, compared to 4% of 28 patients with relaxation therapy. This study used criteria for recovery closer to Matthees’ and colleagues analysis (although in Deale’s study recovery was measured five years after the trial had ended, with a greater chance of recovery with time [10]). Despite the more stringent criteria used, this long term follow up study supports the validity of our finding that 22% of participants were recovered after receiving CBT.

Other studies found similar results. When applying operationalised criteria (no longer fatigued, able to resume activities, and a perception of health and fatigue similar to that of a healthy person) to a cohort of 96 patients who had received CBT, 23% were rated as ‘recovered’ after treatment [12]. Interestingly, this Dutch Group found that having a comorbid illness reduced the numbers recovering, which supports our arguably conservative use of a rating of overall health to assess recovery [12]. Using similar criteria and thresholds
to those used in the PACE trial, Flo and Chalder found that 18% of 140 patients were recovered six months after receiving a course of CBT in a routine clinical setting [13]. These two cohort studies therefore provide further support for our finding that 22% of participants recovered after CBT, rather than the 7% reported by Wilshire et al. [7, 9]. Furthermore, the outcomes from CBT do vary between centres internationally, and it may be that UK outcomes could be improved [14].

**Answers to other criticisms by Wilshire and colleagues**

Wilshire used a study of age appropriate normal data from the six minute walking test (how far a patient can walk in six minutes) to suggest that participants in the PACE trial had not recovered as measured by this test [7]. However, as we have already indicated in journal correspondence, it is not valid to compare the walking test used in PACE to population norms, as it was administered in a non-standard way in the PACE trial [15]. Because of our concerns about the welfare of patients with CFS coping with physical exertion, no encouragement was given to participants during this test. This is in contrast to other studies in which encouragement was used [16], and found to make a difference [17]. In the PACE trial participants were told: “You should walk continuously if possible, but can slow down or stop if you need to.” The PACE walking test data are therefore not comparable with data collected in other studies.

Finally, Wilshire et al believe that the PACE participants’ self-reported outcomes may have been influenced by the expectations they had of the therapy they were allocated to [7]. This is most unlikely however as we measured these expectations before participants started treatment, and found that 72% were confident about APT, compared to 57% about CBT [1]; yet three times as many participants were found to be recovered after CBT than after APT [8].
Summary and conclusions

To summarise, we accept and have made clear in our paper that measuring recovery from CFS with limited data is not an exact science. Consequently the clinical validity of our findings that 22% of patients recovering with either CBT or GET is reasonably disputable [8, 15]. However, a review of the published literature indicates that other studies, including those that used more stringent criteria, report recovery rates closer to the ones we reported than to those reported by Wilshire and colleagues. The important implications for both clinicians and patients are: First, this reanalysis of PACE data does not cast any doubt on the findings that that both CBT and GET are more effective treatments than either APT or SMC alone [1]. Second, whilst definitions of recovery can be disputed the literature supports our conclusions that “...recovery from CFS is possible, and that CBT and GET are the therapies most likely to lead to recovery.” [8].

Conflict of interests

PDW is a member of the Independent Medical Experts Group, a non-department public body, which advises the UK Ministry of Defence regarding the Armed Forces Compensation Scheme. He does paid consultancy for a re-insurance company. MS has received royalties from several publishers of academic books. TC has received royalties from a book regarding self-help for CFS. ALJ and KAG have not conflicts of interest.

Funding

Funding for this trial was received from: UK Medical Research Council, the Department of Health for England, the Department for Work and Pensions, and the Scottish Chief Scientist Office.
References


Likert type scoring on the Chalder fatigue scale. 2016a.


